PRE-SERVICE MATHEMATICS TEACHERS' SUBJECT MATTER KNOWLEDGE OF MEASURES OF CENTRAL TENDENCY

KOMATHI A/P JAIRAMAN

FACULTY OF EDUCATION
UNIVERSITY OF MALAYA
KUALA LUMPUR

2018

PRE-SERVICE MATHEMATICS TEACHERS' SUBJECT MATTER KNOWLEDGE OF MEASURES OF CENTRAL TENDENCY

KOMATHI A/P JAIRAMAN

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

FACULTY OF EDUCATION
UNIVERSITY OF MALAYA
KUALA LUMPUR

2018

UNIVERSITY OF MALAYA
<u>ORIGINAL LITERARY WORK DECLARATION</u>

Name of Candidate: **KOMATHI A/P JAIRAMAN**

Registration/Matric No: **PHA120049**

Name of Degree: **Doctor of Philosophy**

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

PRE-SERVICE MATHEMATICS TEACHERS' SUBJECT MATTER KNOWLEDGE OF MEASURES OF CENTRAL TENDENCY

Field of Study: **MATHEMATICS EDUCATION**

I do solemnly and sincerely declare that:

1) I am the sole author/writer of this Work;
2) This Work is original;
3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
4) I do not have any actual knowledge nor ought I reasonably to know that the making of this work constitutes an infringement of any copyright work;
5) I hereby assign all and every right in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date:

Subscribed and solemnly declared before,

Witness's Signature Date:

Name:

Designation:

Acknowledgment

I would like to express my heartiest gratitude to both my beloved supervisors, Associate Professor Datin Dr Sharifah Norul Akmar Bt. Syed Zamri and Dr Suzieleez Syrene Bt. Abdul Rahim who supervised me throughout my PhD. I am thankful for their patience, sincere guidance, and sharing of vast knowledge in mathematics education with me. Thank you to Dr Leong Kwan Eu, Dr Nabeel Abedalaziz, Mr Norjoharuddeen Bin Mohd Nor, and Ms Santhi Suppiah (Sunway University) for the constructive comments and suggestions given related to my thesis. I also wish to take this opportunity to express my appreciation to my sponsor Kementerian Pelajaran Malaysia for granting me the study leave under the scheme CBBP TB.

I am grateful to the three lectures from both public and private universities who validated the clinical interview tasks and the pre-service mathematics teachers at a public university for their willingness to participate in this study. Thank you to my father, Jairaman Krishnan for his encouragement and not forgetting my siblings, in-laws, and friends who were there for me at my time of need.

A special thank you to my husband, T. Parthipan Rajah and my daughter, Hema Sarita for their support, love, encouragement, patience and sacrifices in many ways which gave me the peace of mind to complete my PhD journey. Finally, my uncountable infinite thanks to my late mother Saraswathy Munusamy. The person who constantly motivated me and kept up with every milestone that I reached in my studies. Although I miss her dearly but I know that it was her blessings that made me reach to this stage. I would like to dedicate my PhD to her.

ABSTRACT

The purpose of this study was to investigate pre-service mathematics teachers' subject matter knowledge of measures of central tendency and their level of subject matter knowledge of measures of central tendency involved in the enhancement of statistical literacy. Data of this study was collected using clinical interview technique based on open-ended questions as instrument. The materials that were collected for the analysis consisted of the audio recordings of the clinical interviews, video recordings of the clinical interviews, subjects' notes in the task sheets, and researcher's notes during the interviews.

A total of fifteen tasks covering four constructs of investigation namely; measures of central tendency with reference to context, measures of central tendency in handling bias, measures of central tendency in problem solving, and measures of central tendency in making inference were devised and used in this study. The subjects of this study consisted of six pre-service mathematics teachers enrolled in a 4-year Bachelor of Science with Education (B. Sc. Ed.) in a public university in Peninsular Malaysia. The selected subjects either majored or minored in mathematics and at the point of data collection, they were either in their third or final year of the program.

The findings of this study were summarised into five main themes of the subject matter knowledge of measures of central tendency that emerged from the clinical interview namely; as averages, data representation, procedures, data summary, and data comparison. The findings related to mean as average, mode as average, median as average, and idea of average were categorised under the theme averages. Mode as data representation, best data representation measure, and robustness of measures were categorised under data representation. Idea of mode, idea of median, backward mean

calculation, representative nature of the mean, and forward mean calculation were categorised under procedures.

Identifying and summarising data using measures of central tendency and summarising equal or unequal sized numerical or graphical data using measures of central tendency were categorised under data summary. Finally, utilising the appropriate measure of central tendency to compare and draw conclusions about equal or unequal sized numerical or graphical data were summarised under data comparison.

The pre-service mathematics teachers' levels of the subject matter knowledge of measures of central tendency in terms of the identified themes were also constructed. Most of the pre-service mathematics teachers attained either high or medium levels. However, it was also found that there were pre-service mathematics teachers who attained low levels for the themes data representation and data comparison,

Regarding the overall level of the subject matter knowledge based on the emerged themes calculated for each of the six pre-service mathematics teachers, only one out of the six pre-service mathematics teachers attained an overall high level meanwhile the remaining five attained a medium level. The findings of this study are hoped to contribute to the discussion on what might constitute the subject matter knowledge of measures of central tendency involved in the enhancement of statistical literacy.

ABSTRAK

PENGETAHUAN SUBJEK TENTANG SUKATAN KECENDERUNGAN MEMUSAT GURU MATEMATIK PRA-PERKHIDMATAN

Tujuan kajian ini adalah untuk mengkaji pengetahuan subjek tentang sukatan kecenderungan memusat dan tahap pengetahuan subjek tentang sukatan kecenderungan memusat yang terlibat dalam pembangunan literasi statistik bagi guru matematik pra-perkhidmatan. Data kajian ini telah dikumpulkan melalui teknik temu bual klinikal berdasarkan soalan-soalan terbuka sebagai instrumen. Bahan-bahan yang dikumpulkan untuk analisis terdiri daripada rakaman audio temu bual klinikal, rakaman video temu bual klinikal, nota subjek dalam lembaran tugas, dan nota penyelidik semasa temu bual.

Sebanyak lima belas tugasan meliputi empat konstruk penyiasatan iaitu: sukatan kecenderungan memusat merujuk kepada konteks, sukatan kecenderungan memusat dalam pengendalian 'bias', sukatan kecenderungan memusat dalam penyelesaian masalah, dan sukatan kecenderungan memusat dalam membuat kesimpulan telah dirancang dan digunakan dalam kajian ini. Subjek kajian ini terdiri daripada enam orang guru matematik pra-perkhidmatan yang sedang mengikuti program 4 tahun Ijazah Sarjana Muda Sains dengan Pendidikan (B. Sc. Ed.) di sebuah universiti awam di Semenanjung Malaysia. Subjek dipilih berdasarkan opsyen matematik utama atau kedua dan semasa fasa pengumpulan data, mereka berada sama ada dalam tahun ketiga atau terakhir program.

Hasil kajian ini telah dirumuskan kepada lima tema utama pengetahuan subjek tentang sukatan kecenderungan memusat yang muncul dari data temuduga klinikal iaitu; purata, perwakilan data, prosedur, ringkasan data, dan perbandingan data. Dapatan kajian merangkumi min sebagai purata, mod sebagai purata, median sebagai

purata, dan idea purata dikategorikan di bawah tema purata. Mod sebagai perwakilan data, sukatan kecenderungan memusat terbaik bagi perwakilan data, dan keteguhan sukatan kecenderungan memusat telah dikategorikan di bawah tema perwakilan data. Idea mod, idea median, pengiraan menggunakan min kebelakang, sifat perwakilan min, dan pengiraan min kehadapan telah dikategorikan di bawah tema prosedur.

Mengenal pasti serta merumuskan data menggunakan sukatan kecenderungan memusat dan meringkaskan data secara berangka atau grafik bersaiz sama atau tidak sama rata menggunakan sukatan kecenderungan memusat telah dikategorikan di bawah tema ringkasan data. Akhir sekali, menggunakan sukatan kecenderungan memusat yang sesuai untuk membanding dan membuat kesimpulan berkaitan data secara berangka atau grafik bersaiz sama atau tidak sama rata dikategorikan di bawah tema perbandingan data.

Tahap pengetahuan subjek tentang sukatan kecenderungan memusat guru matematik pra-perkhidmatan dari segi tema yang dikenalpasti juga telah dibina. Kebanyakan guru matematik pra-perkhidmatan mencapai sama ada tahap tinggi atau sederhana. Walau bagaimanapun, didapati bahawa terdapat guru matematik pra-perkhidmatan yang mencapai tahap rendah untuk tema perwakilan data dan perbandingan data.

Merujuk kepada tahap keseluruhan pengetahuan subjek tentang sukatan kecenderungan memusat berdasarkan tema yang muncul bagi setiap guru matematik pra-perkhidmatan, didapati bahawa hanya satu daripada enam guru matematik pra-perkhidmatan mencapai tahap keseluruhan tinggi manakala lima orang lagi mencapai tahap sederhana. Hasil kajian ini diharap dapat menyumbang kepada perbincangan mengenai apa yang mungkin terkandung dalam pengetahuan subjek tentang sukatan kecenderungan memusat yang terlibat dalam pembangunan literasi statistik.

Table of Contents

Chapter 3 Methodology

List of Figures

List of Tables

List of Appendices

Chapter 1 Introduction

Introduction

This chapter includes the background of the study, statement of problem, research questions, definition of terms, significance of the study, limitations of the study and delimitations of the study.

Background of the study. "Statistics is becoming such a necessary and important area of study" (Ben-Zvi & Garfield, 2008). The incorporation of statistics into most mathematics curriculum has necessitated two main goals for statistics education globally: the preparation of some students for further study of formal statistics and the preparation of all students to use statistics in their daily life (Watson, 2006).

However, studies over the past several decades, have highlighted on the issue that students were unable to relate statistics effectively in their daily lives (Garfield & Ben-Zvi, 2007). As citizens of quantitative world, all school graduates are users of statistics rather than creators and in order for them to function properly as citizens; they must acquire the ability of statistical literacy for decision-making in their lives. Therefore, in the last decade, there had been a rising call to increase the attention on the development of statistical literacy in citizens (Gal, 2002).

Statistical literacy carries many interpretations or definitions. Wallman (1993) interpreted that "statistical literacy is the ability to understand and critically evaluate statistical results that permeate our daily lives-coupled with the ability to appreciate the contributions that statistical thinking can make in public or private, professional and personal decisions" (p. 1).

After almost a decade, Gal (2002) proposed the requirements of statistical literacy in two components. The first component is people's ability to interpret and critically evaluate statistical information, data-related arguments, or stochastic phenomena, which may be encountered in diverse contexts and when relevant. The second component is people's ability to discuss or communicate their reactions to such statistical information which can range from understanding of the meaning, expressing opinions about the implications, or even the concerns regarding the acceptability of given conclusions (Gal, 2002, p. 3).

Meanwhile, Watson (2006) viewed statistical literacy as the meeting point of the statistics curriculum and the daily world. Watson believed that when a statistically literate person comes across statistical information in daily life context, he can make spontaneous decision-making based on his ability to apply statistical tools, general contextual knowledge and critical literacy skills. On the other hand, Garfield and Ben-Zvi (2004) view statistical literacy as a set of skills that students should acquire from any statistics program. They mentioned that the skill to properly evaluate evidence (data) and claims based on data is an important skill that all students should learn as part of their school experiences.

Therefore, the schools had the role in enhancing statistical literacy and produce knowledgeable statistics users. In fact, the role of schools as contributing factor towards literate citizens was acknowledged even way before in 1990s during the emergence of statistics education. This was noted in Frankenstein (1998) which mentioned that the school played a crucial role in enhancing statistical literate students who in turn understand why and how statistics are useful in perceiving and interpreting the world and its complexity.

Several studies in statistics education have described some of the recent developments in statistical literacy. One of the earlier studies was done by Watson (1997). This study identified a framework on the description of three important levels required to develop statistical literacy in students. Later the framework was further developed into a hierarchical construct of six levels of understanding (Watson & Callingham, 2003).

Both the framework and hierarchical construct can be used to assist teachers in the development of statistical literacy by suggesting appropriate aspects of statistical literacy that can be incorporated in the curriculum and lesson planning to ensure the achievement of the highest level of critical thinking.

Several studies have also highlighted on the issue of teachers in the enhancement of statistical literacy (i.e. Watson, 2006; Watson, 2011; Chick & Pierce, 2012). Watson (2006) argued that apart from teachers being the "…big frontier in bringing statistical literacy to all students to prepare them to leave school and enter society" (p. 271), she also firmly believed that teachers should offer their students with productive experiences using real-world examples that demonstrate the utility of statistical concepts involved in the enhancement of statistical literacy.

However, there exists one critical question; whether our teachers have the knowledge to enhance statistical literacy in their classroom. Fennema and Franke (1992) have advocated that, "no one questions the idea that what a teacher know is one of the most important influences on what is done in classroom and ultimately on what students learn" (p. 147).

The National Council of Teachers of Mathematics (NCTM) (2000) *Principles and Standards for School Mathematics* documented that:

> Students learn mathematics through the experiences that teachers provide. Thus, students' understanding of mathematics, their ability to use it to solve problems, and their confidence in, and dispositions toward, mathematics are all shaped by the teaching they encountered in school. The implementation of mathematics education for all students requires effective mathematics teaching in all classrooms. To be effective, teachers must know and understand the mathematics they are teaching and able to draw on that knowledge flexibly in their teaching tasks. (pp. 16-17)

Thus, as mathematics teachers they must be able to draw the knowledge required for the enhancement of statistical literacy flexibly into their teaching. However, there are many knowledge components involved in teaching. Shulman (1986) outlined several strands of teacher knowledge but according to him, the subject matter knowledge which is the content knowledge for teaching had absence of focus in research related to teacher knowledge.

Teachers own subject matter knowledge strongly influences their efforts to help their students learn subject matter (Ball & McDiarmid, 1989). The same applies to statistical literacy. Teachers' subject matter knowledge of statistical topics affected their ability to teach these topics in a way that improves students' statistical literacy (North, Gal, & Zewotir, 2014). Thus, mathematics teachers need to have adequate subject matter knowledge involved in the enhancement of statistical literacy.

On the other hand, it cannot be denied that teachers' subject matter competence undergoes transformation while they are involved in teaching because their initial knowledge of content is enhanced by other knowledge such as knowledge of learners, curriculum, and teaching contexts (Grossman, Wilson, & Shulman, 1989).

Thus, there exists a difference in the subject matter knowledge involved in the enhancement of statistical literacy between in-service and pre-service mathematics teachers. As pre-service mathematics teachers, a solid subject matter knowledge foundation is required upon which greater subject matter competence can be build

(Grossman, Wilson, & Shulman, 1989) which also applies to the subject matter knowledge involved in the enhancement of statistical literacy.

Several statistical topics contribute towards pre-service mathematics teachers' subject matter knowledge involved in the enhancement of statistical literacy. However, measures of central tendency are one of those topics that hold a fundamental position in most mathematics curriculum and also in everyday life. In fact, the concept of mean, median, and mode have been part of the mathematics curriculum even before statistics was emphasized in the curriculum. Thus, pre-service mathematics teachers are not only required to teach these concepts but also to incorporate the enhancement of statistical literacy for the concepts in their teaching.

Nevertheless, studies on teachers which revealed that they lacked in the knowledge of measures of central tendency (Jacobbe & Carvalho, 2011) drew the attention of the present study. Furthermore, studies on teachers which revealed issues such as: they exhibited little or no understanding of the effect of outliers on the mean (Batanero, Godino, & Navas, 1997; Estrada, Batanero, & Fortuny, 2004); applied algorithm without taking account the context; and were unable to invert the mean algorithm (Estrada, Batanero, & Fortuny, 2004; Olfos and Estrella, 2010) indicated serious concerns related to teachers' subject matter knowledge of measures of central tendency namely for the enhancement of statistical literacy.

In the context of Malaysia, measures of central tendency are covered quite extensively in the school mathematics curriculum at various levels. Thus, Malaysian pre-service mathematics teachers need to have adequate foundation related to the subject matter knowledge of measures of central tendency involved in the enhancement of statistical literacy. However, the question is of what should that foundation for the subject matter knowledge of measures of central tendency involved in the enhancement of statistical literacy consist. It is against this background that the present study was initiated.

Statement of Problem

Measures of central tendency are statistical measures which are not only confined to the scientific world. The frequent appearances of these measures in the media and assessment data make them as part of our everyday life. Therefore, these measures also extend to the real world (Jacobbe & Carvalho, 2011; Groth, 2006). The three most common measures of central tendency are the arithmetic mean, the median, and the mode. Each of these measures describes a different indication of the typical or central value in a distribution.

Colloquially measures of central tendency are also called as averages. Accordingly, the three common measures of centre are also the three basic kinds of average. In such situation, the context in which the term average used would indicate which of the three measures of centre is referred to. In scientific, average is normally referred to the arithmetic mean and is hardly used to refer the median or the mode. However, this is unlike in daily life context, where the average can refer to any one of the three measures of centre.

In education, measures of central tendency comprise a fundamental portion of the knowledge that students need to acquire at various levels of education. The education system, has to prepare students with the ideas of these measures not only for further formal studies of these concepts but also for everyday life. However, it is quite saddening to discover that students' experiences involving these measures are often reduced to procedural (e.g. Shaughnessy, 2007, Pollatsek, Lima, & Well, 1981; Mevarech, 1983; Mokros & Russell, 1995; Cai, 1995; Cortina, Saldanha, & Thompson 1999) which questions the preparation of these students for everyday life.

This call us to reflect upon research done on teachers. However, it was found that research done on teachers is still scarce. Studies on teachers revealed that teachers lacked knowledge of this topic and held exaggerated reliance upon procedural knowledge. This led to a disturbing finding that teachers' knowledge of the central tendency concepts might not be far different from their students. (Jacobbe & Carvalho, 2011).

Some of the findings revealed that teachers were ignorant of the relationship between mean, median, and mode; were unable to discern properly that a value which is typical for a given context; and they exhibited little or no understanding of the effect of outliers on the mean (Batanero, Godino, & Navas, 1997; Estrada, Batanero, & Fortuny, 2004). Teachers also carried out routine application of the algorithm without taking account the context; they were unable to invert the mean algorithm (Estrada, Batanero, & Fortuny, 2004; Olfos and Estrella, 2010); and viewed mean as synonymous to average (Leavy & O'Loughlin, 2006). These findings indicate that teachers' might lack in the subject matter knowledge of measures of central tendency for the preparation of students for daily life situations, an important aspect that is emphasized in the development of statistical literacy.

Teachers' subject matter knowledge of statistical topics affect their ability to teach these topics in a way that improves students' statistical literacy (North, Gal, & Zewotir, 2014) which include the topic of measures of central tendency. Although teachers' subject matter knowledge undergoes transformation when they come into teaching because of the enrichment of other knowledge components such as knowledge of learners, curriculum, and teaching contexts but they still need adequate background about their subject prior to their induction to teaching in which greater subject matter competence can be build (Grossman, Wilson, & Shulman, 1989). The same applies to the measures of central tendency.

Therefore, as pre-service mathematics teachers, these teachers should enter the teaching profession with some adequate subject matter knowledge background namely for the topic of measures of central tendency. Making general assumption that all pre-service mathematics teachers who come into the teaching profession are well equipped with the subject matter knowledge might not be true (Even, 1993; Nik Azis, 2003).

Thus, the field should move forward in investigating this issue. Consequently, starting with pre-service mathematics teachers' subject matter knowledge of measures of central tendency involved in the enhancement of statistical literacy is a good point to address the matter since there is a lack of studies related to pre-service mathematics teachers' subject matter knowledge of measures of central tendency involved in the enhancement of statistical literacy.

Moreover, the present study was aimed at the extension on the knowledge related to the development of statistical literacy in which Gal (2002) have highlighted as did not receive much attention over the past one decade. Hence, the present study investigated Malaysian pre-service mathematics teachers' subject matter knowledge of measures of central tendency involved in the enhancement of statistical literacy.

Research Questions

The purpose of this qualitative case study was to investigate pre-service mathematics teachers' subject matter knowledge of measures of central tendency. The present study was also aimed to investigate pre-service mathematics teachers' levels (low, medium, high) of subject matter knowledge of measures of central tendency. In an attempt to investigate the purpose, this study answered the following research questions:

1. What types of subject matter knowledge of measures of central tendency do the pre-service mathematics teachers' have?

2. What levels of subject matter knowledge of measures of central tendency do the pre-service mathematics teachers' exhibit?

Definitions of Terms

The following definitions of terms were used in the present study:

Subject matter knowledge of measures of central tendency - the subject matter knowledge of beginning teachers involves four dimensions: content knowledge, substantive knowledge, syntactic knowledge, and beliefs about the subject matter (Grossman, Wilson, & Shulman, 1989). Shulman (1986) pointed out that teachers' understanding of the subject matter should cover not only knowing "that" but also knowing "why". Shulman's definition of subject matter knowledge was further widened by Ball, Hill, and Bass (2005) as "mathematical knowledge for teaching" which focuses on the content teachers need to know and what ways they need to know this content for teaching. Hence, in this study, the subject matter knowledge of measures of central tendency was conceptualized as the knowledge of the content and organization of the topic measures of central tendency involved in the enhancement of statistical literacy.

Statistical literacy - there are many interpretations given to statistical literacy. Wallman (1993) interpreted that statistical literacy as the ability to understand and critically evaluate statistical information along with the ability to appreciate statistical thinking in various life decisions. Similarly, Gal (2002) elaborated statistical literacy based on two components. The first component is the ability to interpret and critically evaluate statistical information, and the second component is the ability to discuss or communicate their reactions to the statistical information. Statistical literacy is also interpreted as a form of connectivity and skills. For instance, Watson (2006) views statistical literacy as the meeting point of the statistics curriculum and the daily world. Whereas, Garfield and Ben-Zvi (2004) view statistical literacy as a set of skills that students should acquire from any statistics program. However, in the present study, statistical literacy was interpreted as the ability to relate central tendency concepts to real life situations.

Measures of central tendency refer to mean, median, and mode. Colloquially, measures of central tendency are often called averages.

Average in relation to statistical literacy can be referred to either mean, median or mode (Watson, 2006, p. 99).

Mean refers to the sum of all measurements divided by the number of observations in a set of data.

Median refers to the middle value that separates the higher half from the lower half of a set of data.

Mode refers to the most frequent value in a set of data.

Outlier is an observation value which is distant from the other values. Outlier can be due to variability in the measurement or it might indicate experimental error.

Significance of the Study

Past studies on measures of central tendency done on students revealed that students' experiences were very procedural (e.g. Shaughnessy, 2007; Pollatsek, Lima, & Well, 1981; Mevarech, 1983; Mokros & Russell, 1995; Cai, 1995; Cortina, Saldanha, & Thompson, 1999) which highlighted on the need to look at teachers and the knowledge held by them. However, studies done on teachers also revealed the similar and their knowledge related to these measures is not far different from their students (Jacobbe & Carvalho, 2011). This questions on teachers' ability in preparing students for real life situations, an important aspect emphasized in the enhancement of statistical literacy which did not receive much attention as compared to the ideas of measures of central tendency in scientific (Jacobbe & Carvalho, 2011).

Thus, the findings of the present study contributed to the discussion on what might constitute pre-service mathematics teachers' subject matter knowledge of measures of central tendency involved in the enhancement of statistical literacy. The findings of the present study did not only attempt to fill the gap on research and literature related to measures of central tendency on teachers which was found to be scarce (Jacobbe & Carvalho, 2011) but also attempted to fill the gap on understanding the nature of knowledge involved in the development of statistical literacy which Gal (2002) pointed out as lacked in attention more than a decade ago.

Some of the findings of previous studies revealed that teachers were ignorant of the relationship between mean, median, and mode; were unable to discern properly that a value which is typical for a given context; and they exhibited little or no understanding of the effect of outliers on the mean (Batanero, Godino, & Navas, 1997; Estrada, Batanero, & Fortuny, 2004). Teachers also carried out routine application of the algorithm without taking account the context; they were unable to invert the mean algorithm (Estrada, Batanero, & Fortuny, 2004; Olfos and Estrella, 2010); and viewed mean as synonymous to average (Leavy & O'Loughlin, 2006). These findings indicated that teachers' might lack in the subject matter knowledge of measures of central tendency involved in the enhancement of statistical literacy.

The findings of the present study have specifically identified through the emerged themes that there are five types of subject matter knowledge of measures of central tendency needed for the enhancement of statistical literacy namely; averages, data representation, procedures, data summary, and data comparison. Through the five emerged themes, some of the pre-service mathematics teachers in the present study had shown better level of subject matter knowledge in certain themes such as procedures and data summary. Whereas, in the other themes like averages, data representation and data comparison, they have shown a lower level of subject matter knowledge of measures of central tendency. The pre-service mathematics teachers in this study had no issues inverting the mean algorithm or understanding the effect of outliers on the measures but they exhibited issues related to the term average or discern properly that a value which is typical for a given context.

Thus, the findings of the present study contribute to which knowledge the pre-service mathematics teachers lacked in. It is hoped that these findings inform teacher educators, teacher trainers, and education policy makers about the capacity of our Pre-service Mathematics Teacher Education and Teacher Preparation programs in cultivating the necessary subject matter knowledge of measures of central tendency for the enhancement of statistical literacy. These programs can be revisited or reconstructed with the ideas of measures of central tendency in the five emerged themes with hope to prepare future teachers with adequate subject matter knowledge of measures of central tendency.

Apart from this, the findings of the present study suggest that In-service Teacher Education programs that are organised and conducted by the State Education Departments, Teacher Training Institutes, or Curriculum Development Centre to facilitate current mathematics teachers' revisitation on the topic of measures of central tendency and incorporate ideas of measures of central tendency in the five emerged themes in hope to train the in-service teachers with the ideas of measures of central tendency important for the enhancement of statistical literacy.

The findings of the present study would also suggest a revisitation of the mathematics curriculum at various levels of school education. The findings uncovered aspects like pre-service mathematics teachers' lack of knowledge of the term average, related the term average to the mean, and were unaware that average can be referred to any of the three measures of central tendency given to the context of everyday life. Moreover, they also lacked in the knowledge of measures of central tendency in various representations. Thus, the curriculum should state clearly the terms of measures of central tendency given in different contexts and suggest examples on activities or tasks of various representations that can be incorporated in the lesson planning and teaching.

Finally, the findings of the present study would also like to suggest that the school text books or references to incorporate activities that go beyond calculation and procedures of measures of central tendency. The text book activities should incorporate the ideas of measures of central tendency in the five emerged themes using everyday examples.

Limitations of the Study

The present study involved only six pre-service mathematics teachers. The participants were drawn from the pre-service mathematics teachers who enrolled in a 4-year Bachelor of Science with Education (B.Sc.Ed.) program in a public university in Peninsula Malaysia. This purposive sampling is considered to be small to make any logical and reasonable claims.

The findings of the present study could not be generalized to other pre-service mathematics teachers enrolled in the same program or in other education programs in the selected public university where the data was collected. The findings of the present study also could not be generalized to other pre-service mathematics teachers in other universities and teacher training institutes. Thus, the limitation of the present study dealt with generalizability.

The present study looked at only ideas of measures of central tendency involved in the enhancement of statistical literacy and did not examine other ideas or properties of measures of central tendency. The present study also did not include the connections between the ideas of measures of central tendency with other statistical ideas such as measures of dispersion. Hence, the findings of the present study were confined to only ideas of measures of central tendency involved in the enhancement of statistical literacy.

The present study employed the case study research design in order to investigate in-depth, pre-service mathematics teachers' subject matter knowledge of measures of central tendency. However, some of the characteristics of the case study design do pose certain limitations which also applied to the present study such as the product to be very lengthy (Merriam, 1998; Yin, 2003).

Clinical interview techniques were used to collect data in which the technique itself is not free of limitations which also applied to the present study. The clinical interview technique has been criticized for (a) "lack of standardization of procedures, and (b) lack of information for precise replication" (Nik Azis, 1987, p. 61). Nik Azis (1987) mentioned that "the technique raised some methodological questions in some circles, because reflection, contingency, and creativity are considered to be unorthodox from the perspective of the standardized, objective, replicable scientific method" (p. 61).

Since, the researcher was the primary instrument involved in the data collection and analysis in this qualitative case study, the entire process was bounded by the integrity and also the sensitivity of the researcher (Merriam, 1998). The researcher made the inferences based on the participants' verbal and nonverbal behaviours.

Delimitations of the Study

The present study is delimited to the pre-service mathematics teachers who are enrolled in a 4-year Bachelor of Science with Education (BSc. Ed.) program in a selected public university in Peninsular Malaysia who either majored or minored in Mathematics and were either in their third or final year of the program. The researcher had defined the boundary for the purposive sample selection of the study to the above criteria.

The researcher had also defined the boundary for the investigation based on the two research questions; what types of subject of subject matter knowledge of measures of central tendency do pre-service mathematics teachers' have? And what levels of subject matter knowledge of measures of central tendency do pre-service mathematics teachers' exhibit?

The present study is delimited to the Gal's statistical knowledge of measures of central tendency which focused at measures of central tendency and data, familiarity with basic terms and ideas related to measures of central tendency, and conclusions or inferences reached based on the measures of central tendency. Furthermore, the clinical interview tasks were designed based on the four constructs namely; with reference to the context, in handling bias, in problem solving, and in making inferences. The scope of the investigation of the subject matter knowledge of measures of central tendency was on the aspects involved in the enhancement of statistical literacy.

Chapter 2 Review of the Literature

Introduction

This chapter discusses on the literature related to the central idea of the present study; teachers' subject matter knowledge of measures of central tendency. The included sections in this chapter are: the theoretical perspective, the conceptual framework, subject matter knowledge, research related to subject matter knowledge, levels of subject matter knowledge, research related to Malaysian pre-service teachers' subject matter knowledge, research related to students' understanding of measures of central tendency, research related to teachers' understanding of measures of central tendency, measures of central tendency in the Malaysian mathematics curriculum, the nature of statistical literacy, ideas of measures of central tendency involved in the enhancement of statistical literacy, and Learning Mathematics for Teaching: The Coding Rubric. Each section under the literature review was concluded with a summary of the section.

Theoretical Perspective of the Study

Recent research on teacher knowledge focused on exploring the nature, form, organization, and content of teacher knowledge instead of trying to find the relationship between teacher knowledge and student achievement (Grossman, Wilson, & Shulman, 1989). Hence, the present study was also focused in the same direction; exploring the nature of teacher knowledge.

A perspective on teacher knowledge was introduced by Shulman (1986). Shulman offered a set of hypothetical domains of teacher knowledge in teaching which distinguishes three types of content knowledge in teaching; subject matter content knowledge, pedagogical content knowledge, and curricular knowledge.

Shulman referred the subject matter knowledge as the content knowledge for teaching that grows in the mind of teachers. Pedagogical content knowledge goes into the dimension of subject matter knowledge for teaching. While, the curriculum knowledge is referred to the knowledge of the programs designed for the teaching of particular subjects at given levels (Shulman, 1986, p. 9-10).

However, among the three types of teacher knowledge, Shulman mentioned that the subject matter knowledge had absence of focus in the research on teaching and referred to as the "missing paradigm" which might have serious consequences both for policy and for research (Shulman, 1986, p. 7).

Different subject matter areas will have different ways of discussing the content structure of knowledge. The content knowledge goes beyond knowledge of facts or concepts of a domain. It requires understanding the structures of a subject which include both the substantive and syntactic structures (Shulman, 1986, p. 9). Beliefs about the content also influence the subject matter especially in what and how teachers choose to teach. Apart from this, teachers' beliefs also influence their 'orientation' towards the subject matter (Grossman, Wilson, & Shulman, 1989).

The same goes to the subject matter knowledge of measures of central tendency. The subject matter knowledge of measures of central tendency requires beyond knowledge of facts or concepts of central tendency. It requires understanding substantive structures which includes the variety of ways in which the basic concepts or principles of measures of central tendency are organized to incorporate its facts. It also requires syntactic structures which are ways in which truth or falsehood, validity or invalidity are established based on measures of central tendency. Apart from this, certain beliefs regarding the content of measures of central tendency also influence its related justifications.

On the other hand, a teacher might not have the same foundation in regards to the subject matter knowledge of measure of central tendency. Although, a teacher learns more about the content of measure of central tendency when he or she is involved in the preparation and teaching of this topic but as a pre-service mathematics teacher, he or she does come into the teaching profession with some background on the subject matter knowledge that he or she is going to teach. This background provides the foundation for the accumulation of greater subject matter competence (Grossman, Wilson, & Shulman, 1989).

A well-prepared pre-service mathematics teacher who possesses content knowledge of measures of central tendency knows the central concepts of measures of central tendency and all the organizing principles. He or she will be aware of the relationship between the procedures with the conceptual issues related to measures of central tendency.

When a pre-service mathematics teacher lacks the content knowledge of measures of central tendency, he or she might not be aware that the idea of average can involve all three measures of central tendency or certain aspects involved in the presentation of the average value will indicate which of the three measures is being referred to.

For example, if a real-world situation involving the number of children of a family and the presentation of the average value for the number of children is in the form of a decimal number with more than two decimal places then the average here is more likely referring to the mean. On the other hand, if an average is obtained based on the percentage than most likely the average is based on the mode and if it involves data on human population then the average most likely is based on the median.

The content knowledge of measures of central tendency does not exist independent of the deeper structures of the discipline; both the substantive and syntactic. A well-prepared pre-service mathematics teacher in terms of the substantive structures will recognize a data set and all the other alternative forms that can represent an average value.

He or she will understand why in certain situation the median is more suitable than the mean in summarising a data set. He or she will not only be able to calculate the mean or median but will also have a sense of the statistical meaning underlying the meaning of the mean or median.

A well-prepared pre-service mathematics teacher in terms of syntactic structures will understand that measures of central tendency are not just measures that are calculated using algorithm but are measures that explain about a data set.

He or she will not only be capable of defining or explaining facts or concepts related to measures of central tendency but will be able to explain the claims made related to these measures; when competing claims are offered, how the controversy be adjudicated.

Furthermore, a pre-service mathematics teacher need not only understand that a fact or concept related to measures of central tendency is so but further understand why it is and on what grounds its warrant can be asserted.

A pre-service mathematics teacher also holds certain beliefs regarding the content that he or she will be teaching which might influence what and how he or she choose to teach. If a pre-service mathematics teacher believes that measures of central tendency are just algorithms than he or she will choose to teach the concepts in this manner.

Apart from this, a pre-service mathematics teacher might hold to a belief that an average is synonym to the mean or a belief that only mean can be used to summarise a data set which will have an influence on his or her teaching.

Thus, the foundation on the subject matter knowledge of measures of central tendency in our pre-service mathematics teachers actually prepares them with the content knowledge for teaching the ideas of central tendency both for further formal studies and for daily life uses.

Nevertheless, the present study sought to investigate pre-service mathematics teachers' subject matter knowledge of measures of central tendency related to daily life; an aspect greatly emphasized in the enhancement of statistical literacy. Therefore, although the present study in general followed Shulman's perspective on the teachers' subject matter knowledge but the focus of the subject matter knowledge investigation was based on Gal's (2002) model of statistical literacy because Shulman's perspective on teachers' subject matter knowledge underpin the theoretical foundation behind Gal's model of statistical literacy which outlined the subject matter knowledge needed for the development of statistical literacy.

Gal (2002) through his proposed model of statistical literacy outlined the knowledge bases and other enabling processes that should be available in adults and by implication to school graduates to function as statistically literate individuals. These components are seen to illustrate the subject matter knowledge involved in the enhancement of statistical literacy. Thus, the present study centered the investigation on the subject matter knowledge of measures of central tendency involved in the enhancement of statistical literacy based on Gal's model of statistical literacy.

Conceptual Framework of the Study

The conceptual framework of the present study was based on the literature review and the theoretical perspective presented in the preceding section. The diagram in Figure 2.1 illustrates the conceptual framework of the present study**.** The present study was aimed at investigating the subject matter knowledge of measures of central tendency involved in the enhancement of statistical literacy. Therefore, the study focused on the model of statistical literacy presented by Gal (2002).

The knowledge component of Gal's model consists of five cognitive elements: literacy skills, statistical knowledge, mathematical knowledge, context knowledge and critical questions whereas the dispositional component consists of two elements: critical stance and belief and attitudes (elaborated further in the literature review).

The proposed model is viewed as a context-dependent, dynamic set of knowledge and dispositions that together enable statistically literate behavior (Gal, 2002, p. 4). The five elements of the knowledge component form the base and jointly contribute to a teacher's ability to comprehend, interpret, critically evaluate and if required to react to statistical messages in various forms such as oral, written text, tables or graphical representations (Gal, 2002).

However, in the present study, the investigation on the subject matter knowledge of measures of central tendency was focused on the statistical knowledge of measures of central tendency which included the connection between measures of central tendency and data, familiarity with basic terms, and ideas related to measures of central tendency and conclusions reached based on measures of central tendency (Gal, 2002, p. 10).

Pre-service mathematics teachers should possess some knowledge regarding the origins of the data that was used in any reporting involving measures of central tendency. They must also have the knowledge on how the data is produced, and be aware of the contribution of a good design for data production to the possibility of answering specific questions (Cobb & Moore, 1997).

Pre-service mathematics teachers must also understand that the need to reduce data is to identify key features and trends in spite of noise and variation because this provides the basis for accepting the use of measures of central tendency as statistical summaries. Measures of central tendency act as tools for relating information from data producers to data consumers (Wild & Pfannkuch, 1999).

Measures of central tendency are noted for its centrality in descriptive statistics. Gal (1995) mentioned that consumers (in the context of the present study refers to pre-service mathematics teachers) of statistical reports should know that means and medians are simple ways to summarise a data set. These measures show the centre of the data. However, the mean is affected by extreme values more so than the median. The ideas of measures of central tendency can mislead when the distribution or shape of the data on which they are based is very uneven or bimodal. Measures of central tendency can also mislead if these measures are calculated based on data or sample that is not representative of the whole population under study.

Measures of central tendency are typical ways of summarising data. However, due to the existence of different designs for collecting data, and that sampling processes involved, pre-service mathematics teachers need to have some knowledge in regards to how data are analysed, conclusions are obtained and the related problems occurred. Pre-service mathematics teachers need to be sensitive to the possibility of biases in inference and puts forward their concern about the stability and generality of findings (Gal, 2002, p. 13).

Based on the above interpretations of the ideas of measures of central tendency involved in the enhancement of statistical literacy. The investigation of the subject matter knowledge of measures of central tendency involved in the enhancement of statistical literacy was based on the following four main constructs in which Watson (2006) highlighted as important for the development of statistical literacy.

Measures of central tendency with reference to context – The idea of average in statistical literacy can be associated with all the three measures of central tendency (Watson, 2006, p. 99). Assumption that the arithmetic mean is the average all the time is not correct. Being able to relate algorithm to concrete situations such as in the presentation of the average to two decimal places or the idea of mean which does not preserve numerical set (Batanero, Cobo, & Diaz, 2003; Watson, 2006) are considered important in relation to statistical literacy.

Apart from this, knowing typical behaviour when percentage is used should indicate that the average is referring to the mode. Reports based on human population should give rise to the awareness of the shape of the distribution and the presence of outliers. The averages in such reports might indicate that the median might have been used (Watson, 2006, p. 99).

Relating the average with the associated data set from which it is constructed also reflects statistical literacy (Mokros & Russell, 1995). Statistical literate readers of data when they see an average value; they think of the different distributions that these indicators could represent.

Making sensible decisions on the possible data set and how average can be obtained from the data set will provide some valuable insights on the representativeness of average and the appreciation of average in reducing the complexities of the data. Representativeness of average as a typical behaviour, "middle value", a balance point, a fair share or an algorithm (Mokros & Russell, 1995; Strauss & Bichler, 1988; Watson, 2006) are all considered important when statistical literacy is concerned.

Measures of central tendency in handling bias - Another important link to the ideas associated to measures of central tendency in relation to statistical literacy is the discussion on extreme values and which of the three measures of centre is least likely to be biased by these extreme values (Watson, 2006, p. 110). Understanding the distinctive features of the arithmetic mean and median can lead to an appreciation on why it may be important to use one or the other (Watson, 2006, p. 110).

Measures of central tendency in problem-solving - Deconstructing of the mean can lead to a consideration of variation in the data set it represents. The representative nature of the mean can be considered in making sense of why data sets need to be combined to obtain a weighted mean. Although, algorithms are necessary but understanding the connections of the algorithms with the purposes of using summary statistics is equally important (Watson, 2006, p. 113). In order to solve weighted mean problems, the knowledge of applying the algorithm for the arithmetic mean in both the "forward" and "backward" direction is necessary (Watson, 2006, p. 114).

Measures of central tendency in making inference - Measures of central tendency in reducing the data to single value(s), leads to inference stage of statistical investigations. One of the goals of statistical literacy is recognising measures of central tendency as significant tools for comparisons or judgements about data sets without being specifically told of its use (Watson, 2006, pp. 118-119). In addition to this, average as "signal in noise" takes on the perspective that each observation is an estimate of the unknown but specific value (Konold & Pollatsek, 2004). The mean should be considered as a suitable measure and should be viewed as a meaningful real-world concept (Leavy & O'Loughlin, 2006, Pollatsek, Lima & Well, 1981) and not just a calculated measure. In any reporting of average, questions concerning range, outliers, spread, and shape should arise in the mind of any statistically literate individuals (Watson, 2006).

Therefore, in order to understand the nature of subject matter knowledge of measures of central tendency held by our pre-service mathematics teachers, the researcher guided her investigation based on the above theoretical and conceptual focus.

Apart from this, the present study also determined pre-service mathematics teachers' levels (low. medium, high) of subject matter knowledge of measures of central tendency using the coding rubrics (see Appendix I) adapted from the Learning Mathematics for Teaching (LMT) (2006) project by Hill, Ball, Bass and Schilling.

In order to understand a pre-service mathematics teacher's level of subject matter knowledge of measures of central tendency for each of the four constructs, the coding process adapted from Learning Mathematics for Teaching (LMT) (2006) project by Hill, Ball, Bass, and Schilling was used.

The following elaborates on the coding process. The researcher determined whether the statistical element was present (P) or not present (NP). If the statistical element was present (P), then she marked: (a) appropriate (A) if the pre-service mathematics teacher's use of the statistical element was appropriate, accurate, or correct; or marked (b) inappropriate (I) if the pre-service mathematics teacher's use of the statistical element was inappropriate, inaccurate, or incorrect. If the statistical element was not present (NP), then she marked (a) appropriate (A) if the absence of the statistical element seemed appropriate or not problematic; or marked (b) inappropriate (I) if the absence of the statistical element seemed inappropriate or problematic (i.e., the statistical element should have been present) (adapted from LMT, 2006). The following diagram illustrates the coding thought process. (Further elaborated in Learning Mathematics for Teaching: The Coding Rubric section in the Review of the Literature)

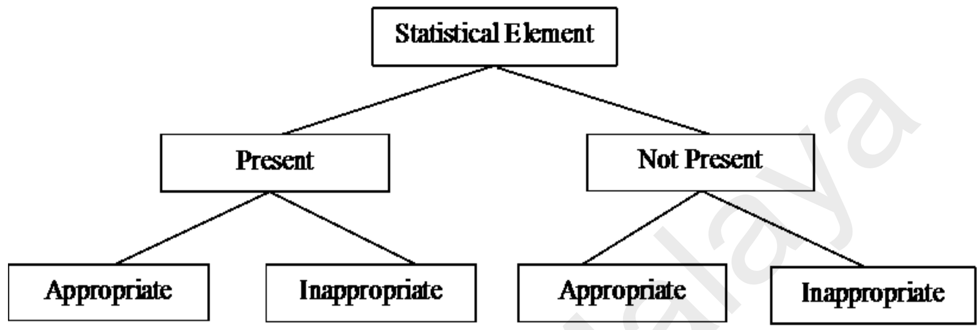


Figure 2.1. Coding thought process

A preservice mathematics teacher's level (low, medium, high) of subject matter knowledge of measures of central tendency for each of the construct was determined based on the percentage of appropriate statistical elements of the subject matter knowledge of measures of central tendency. The following illustrates the computation of the percentage of appropriate statistical elements of subject matter knowledge of measures of central tendency:

Percentage of appropriate statistical elements of subject matter knowledge of measures of central tendency was calculated based on the following formula = $\frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)}$ x 100%, where nPA, nPI, nNPA, and nNPI represented the number of codes that are coded as "present and appropriate (PA)", "present and inappropriate (PI)", "not present and appropriate (NPA)", and "not present and inappropriate (NPI)", respectively.

A preservice mathematics teacher's level (low, medium, high) of the overall subject matter knowledge of measures of central tendency was determined using the grading system used in the university where the data of the study was collected. In order to obtain a preservice mathematics teacher's level of subject matter knowledge of measures of central tendency, first the mean percentage of the subject matter knowledge for all four constructs namely; with reference to the context, in handling bias, in problem solving, and in making inference was calculated.

The following describes on how the mean percentage was calculated for all four constructs. Mean percentage of the subject matter knowledge of measures of central = $\frac{A+B+C+D}{4}$ % where A represented the percentage of subject matter knowledge with reference to context, B represented the percentage of subject matter knowledge in handling bias, C represented the percentage of subject matter knowledge in problem solving, and D represented the percentage of subject matter knowledge in making inference.

If a preservice mathematics teacher achieved 80% and above as the mean percentage of the overall subject matter knowledge, he or she was assigned a high level of subject matter knowledge of measures of central tendency. If a preservice mathematics teacher's achievement was 40% to less than 80%, he or she was assigned a medium level of subject matter knowledge of measures of central tendency. Finally, if a preservice mathematics teacher's achievement was less than 40%, he or she was assigned a low level of subject matter knowledge of measures of central tendency.

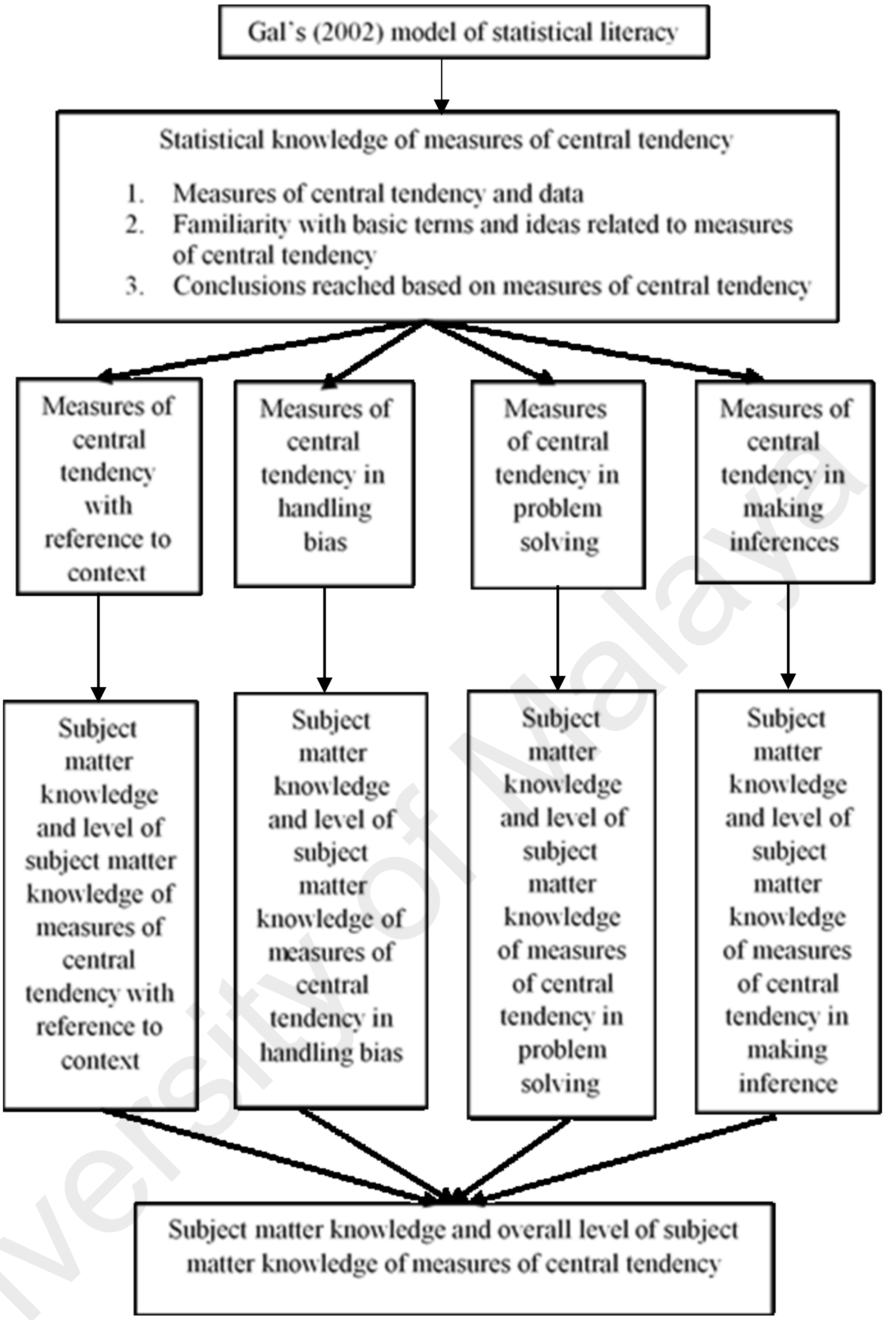


Figure 2.2. Conceptual framework of the study

Subject Matter Knowledge

Teaching is an activity of both thought and action. Teachers decide on the content and goals, materials and approaches, in order to help different students, learn a variety of concepts and skills. It was highlighted that future research should focus on the knowledge teachers need to teach statistics (Shaughnessy, 2007). However, there are many components of teacher knowledge that are required of a teacher and the subject matter knowledge is one of those.

"Philosophical arguments as well as common sense support the conviction that teachers' own subject matter knowledge influences their efforts to help students learn subject matter" (Ball & McDiarmid, 1989).

Even (1990) mentioned that:

> The teacher's role is to help the learner achieve understanding of the subject matter. But in order to do so the teachers themselves need to have solid knowledge of the subject matter. A teacher who has solid mathematical knowledge for teaching is more capable of helping his/her students achieve a meaningful understanding of the subject matter. (p. 521)

The subject matter knowledge of beginning teachers involves four dimensions: content knowledge, substantive knowledge, syntactic knowledge, and beliefs about the subject matter (Grossman, Wilson, & Shulman, 1989).

Content knowledge is referred to the "stuff" of a discipline that includes the factual information, organizing principles and central concepts (Grossman, Wilson, & Shulman, 1989, p.27). Teachers' lack of content knowledge can affect how they critique textbooks, how they select material to teach, how they structure their courses, and how they conduct instruction (Grossman, Wilson, & Shulman, 1989, p. 28).

Substantive knowledge is referred to substantive structures of a discipline which include the explanatory frameworks or paradigms that are used both to guide inquiry in the field and to make sense of data (Schwab, 1978 in Grossman, Wilson, & Shulman, 1989, p. 29). "The substantive structures are the variety of ways in which the basic concepts and principles of the discipline are organized to incorporate its facts" (Shulman, 1986, p. 9).

Ball (1990) mentioned that substantive knowledge does not only refer to teachers being able to do mathematics to produce correct answers but also refer to teachers having a sense of the mathematical meanings underlying the concepts and processes. Substantive knowledge is not just a collection of disparate facts and procedures but should be a collection of interconnected concepts and procedures.

Syntactic knowledge is referred to the syntactic structures of a discipline which are the canons of evidence that are used by members of the disciplinary community to guide inquiry in the field. They are the means by which new knowledge is introduced and accepted into that community (Schwab, 1978 in Grossman, Wilson, & Shulman, 1989, p. 30).

According to Shulman (1986, p. 9), the syntactic structure of a discipline is the set of ways in which truth or falsehood, validity or invalidity, are established". Syntax of a discipline provides the rules for determining which claim has greater warrant, when competing claims exist. It is a set of rules for determining what is legitimate to say in a disciplinary domain and what goes against the rules. A lack of syntactic knowledge may seriously limit teachers' abilities to learn new information in their fields.

Teachers also hold certain orientation toward the subject that they teach and also hold certain belief related to the content that they teach (Grossman, Wilson, & Shulman, 1989, p. 31). Teachers' understanding of subject matter does not only involve the intellectual aspects but also the essence of the subject matter. Thus, teachers' intellectual resources and dispositions largely determine their capacity to engage students' minds and hearts in learning (Ball & McDiarmid, 1989, p. 5).

However, it is regarded that the subject matter knowledge of experienced teachers is different than of beginning teachers. Teachers build upon their knowledge of subjects when they prepare to teach. Mathematics teachers' subject matter knowledge undergoes transformation while they are involved in teaching. Their initial knowledge of content is enhanced by other knowledge such as knowledge of learners, curriculum, and teaching contexts (Grossman, Wilson, & Shulman, 1989).

While mathematics teachers learn more about the subject matter when they come into teaching, they also require adequate background about their subject prior to their induction to teaching. Therefore, as pre-service mathematics teachers, these teachers should have a solid foundation of the subject matter knowledge upon which greater subject matter competence can be build (Grossman, Wilson, & Shulman, 1989).

Shulman (1986) also pointed out teachers' understanding of the subject matter should cover knowing "that" and knowing "why". Shulman mentioned that:

> We expect that subject-matter content understanding of the teacher be at least equal to that of his or her lay colleague, the mere subject-matter major. The teacher need not only understand *that* something is so; the teacher must further understand *why* it is so. (Shulman, 1986, p. 9)

However, Shulman (1986, p. 7) highlighted that the subject matter knowledge had absence of focus in the research on teaching and referred to as the "missing paradigm" which might have serious consequences both for policy and for research.

One study that explored further on Shulman's subject matter knowledge and widened the definition on "mathematical knowledge for teaching" was done by Ball, Hill, and Bass (2005). They seek to find what content teachers need to know to teach and what ways they need to know the content to use in their teaching.

The focus was on the extra knowledge of mathematics required in teaching because teaching mathematics is not merely doing problems as students watched but it involves explaining, listening and examining students' work. It also involves making choices on the appropriate and useful models and examples.

Being able to do all these demands additional mathematical insight and knowledge. Ball and his colleagues believed that defining mathematical knowledge for teaching specifically in this way will provide a basis for what teachers are taught and increases the possibilities of teachers utilising what they are taught when they teach. They hypothesized teachers' opportunities to learn mathematics for teaching could be better designed if the types of knowledge required are identified clearly.

Ball et al. (2005) proposed a refinement of Shulman's (1986) categories of teacher knowledge as shown in Figure 2.1 to support their hypotheses that showed the correspondence between their current map of the domain of content knowledge for teaching and Shulman's (1986) initial categories: subject matter knowledge and pedagogical content knowledge.

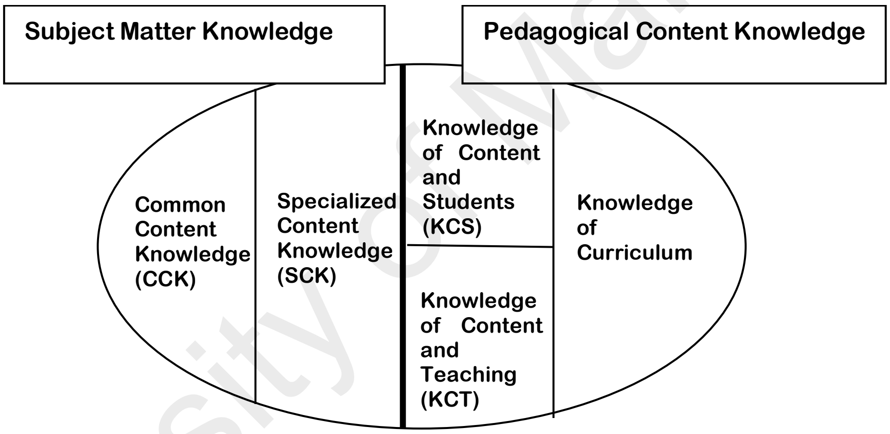


Figure 2.3. Shulman's Original Category Scheme (1985) compared to Ball, Hill, and Bass (2005)

Shulman' s category of content knowledge was subdivided into common content knowledge and specialized content knowledge whereas the pedagogical content knowledge was subdivided into knowledge of content and students, knowledge of content and teaching, and knowledge of curriculum.

Ball et al. (2005) highlighted on a special kind of subject matter knowledge which is the specialized content knowledge which was found evident by Hill, Schilling, and Ball (2004) in their study. "Specialized content knowledge is mathematical knowledge beyond that expected of any well-educated adult but not yet requiring knowledge of students or knowledge of teaching" (Ball, Thames, Phelps, 2008, p. 9).

Teachers need to know a body of mathematics not typically taught to students which is also different from what Shulman described under the notion of pedagogical content knowledge. Hill, Ball, and Schilling (2008) proposed model of mathematical knowledge for teaching (MKT), extending from the earlier refinement of Shulman's (1986) teacher knowledge categories proposed Ball, Hill, and Bass (2005) shown in Figure 2.2 with an additional category: Knowledge at the Mathematical Horizon.

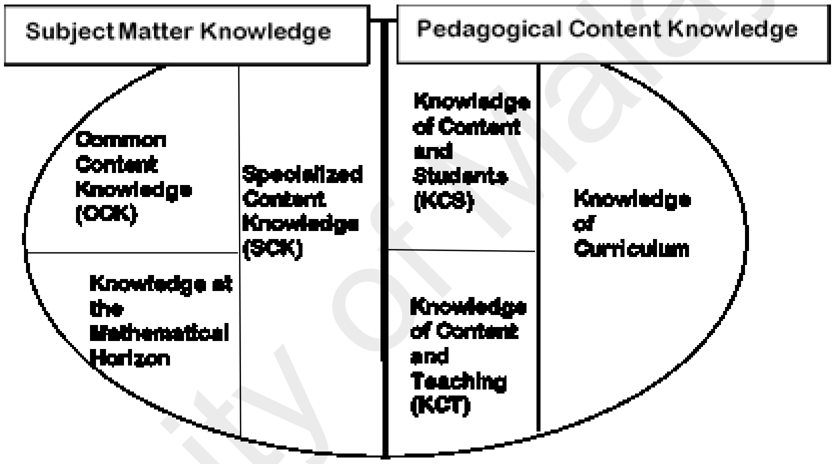


Figure 2.4. Mathematical Knowledge for Teaching (MKT) (Hill & Ball, 2009, p. 70)

Although, the interpretations and names given to the knowledge domains differ from Shulman's original notion of teacher knowledge but six main categories of MKT were distinguished. Category 1 is what Shulman likely meant by his original subject matter knowledge. Category 2 is a newer conceptualization entail from the earlier work of Ball and his colleagues (Ball, Hill, & Bass, 2005; Ball, Lubienski, & Mewborn, 2001; Hill, Ball, & Schilling, 2008; Hill, Schilling, & Ball 2004). Categories 4 to 6 reflect upon Shulman's pedagogical content knowledge.

1. Common content knowledge (CCK) involves the mathematical knowledge and skills used in general settings that teachers should possess which need not be unique to teaching (Hill, Ball, & Schilling, 2008, pp. 377-378).

2. Specialized content knowledge (SCK) involves the mathematical knowledge and skills teachers should possess which they use in teaching but is still independent of the learners or teaching (Hill, Ball, & Schilling, 2008, p. 378).
3. Knowledge at the mathematical horizon involves the understanding of the broader set of mathematical ideas to which a particular idea connects (Hill & Ball, 2009).
4. Knowledge of content and students (KCS) that involves the amalgamated knowledge about students and mathematics (Hill, Ball, & Schilling, 2008, p. 378).
5. Knowledge of content and teaching (KCT) that involves in the design of instruction such as how the examples or representations are chosen, and how accurate mathematical ideas are guided in student discussions (Thames, Sleep, Bass, & Ball, 2008, p. 5).
6. Knowledge of curriculum involves the ways to sequence and structure the development of mathematical topic.

Ball, Thames and Phelps (2008) further listed some of the mathematical tasks for teaching that entail the specialized knowledge:

a. Presenting mathematical ideas
b. Responding to students "why" questions
c. Finding an example to make a specific mathematical point
d. Recognizing what is involved in using a particular representation
e. Linking representations to underlying ideas and to other representations
f. Connecting a topic being taught to topics from prior or future years
g. Explaining mathematical goals and purposes to parents
h. Appraising and adapting the mathematical content of textbooks
i. Modifying the task to be either easier or harder
j. Evaluating the plausibility of students' claims (often quickly)
k. Giving and evaluating mathematical explanations
l. Choosing and developing useable definitions
m. Using mathematical notation and language and critiquing its use
n. Asking productive mathematical questions
o. Selecting representations for particular purposes
p. Inspecting equivalencies (p. 10)

Even though, each of the listed mathematical tasks are something that teachers routinely do but if taken together, these tasks make for some rather unique mathematical requirements for teaching. The listed tasks of teaching depend on mathematical knowledge and are relatively independent of knowledge of students or of teaching. These are tasks that require knowing how knowledge is generated and structured in the discipline (Ball, Thames, & Phelps, 2008, p. 12).

Most of the above-mentioned studies focused on the general idea of the subject matter knowledge in mathematics. Thus, giving an emergence to research on teachers' mathematics topic specific subject matter knowledge because analysing teachers' subject matter knowledge in general will not inform what subject matter knowledge teachers need to have to teach a specific mathematics topic (Even, 1990, p. 522).

"Teachers' subject matter knowledge about a specific mathematical topic is influenced by what they know across different domains of knowledge" (Even, 1990, p. 523). Even identified an analytic framework of subject matter knowledge that highlighted seven aspects which form the main facets of teachers' subject matter knowledge about a specific mathematical topic:

Essential features
Different representations
Alternative ways of approaching
The strength of the concept
Basic repertoire
Knowledge and understanding of a concept
Knowledge about mathematics (Even, 1990, pp. 523-527)

Essential features deal with the concept image with attention to the essence of the concept. Vinner (1983) defines concept image as the set of all 'pictures' that have ever been associated with the concept together with the set of properties associated with the concept in the person's mind.

Different representations refer to the understanding of a concept from one representation to another representation because understanding a concept in one representation might differ from the understanding in another representation. Therefore, teachers need to understand concepts in different representations and should be able to translate and form linkages among and between the concepts (Even, 1990, p. 524).

Alternative ways of approaching – "The appearance of a complex concept in various forms, representations, labels and notations are enhanced by the different uses of the concept in the different divisions of mathematics, other disciplines or everyday life" (Even, 1990, p. 525). In situations when more than one approach can be used, some approaches are more suitable than other. Thus, decisions on the best approach between the available one need to be made. Teachers must know the main alternative approaches and their uses.

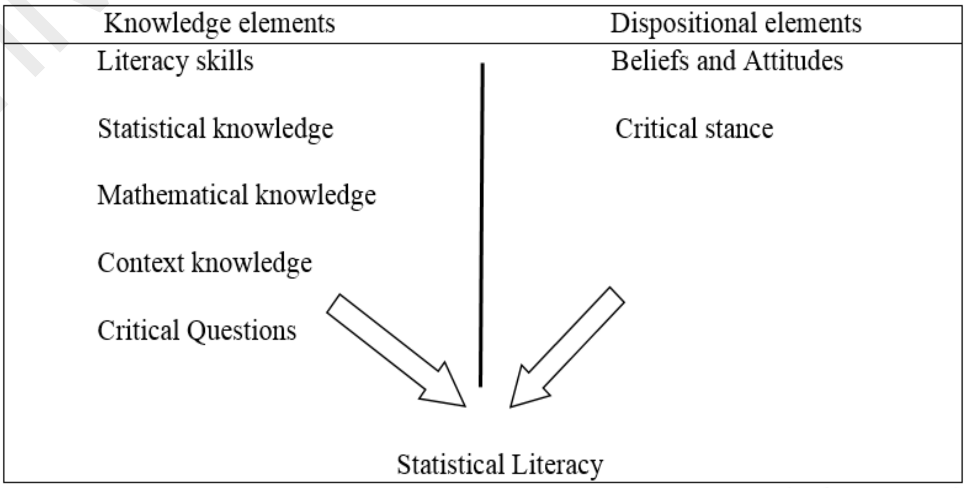
The strength of the concept – "Concepts become important and powerful because there is something special about them which is very unique and opens new possibilities" (Even, 1990, p. 525). Therefore, teachers need to have a good understanding of the unique powerful features of the concept.

Basic repertoire – "The basic repertoire includes powerful examples that illustrate important principles, properties, theorems, etc." (Even, 1990, p. 525). Every mathematical topic or concept should have specific examples that are well known and familiar so that these examples are readily available for use. However, meaningful acquisition of basic repertoire with deep understanding will allow it to be used appropriately and wisely.

Knowledge and understanding of a concept should include both conceptual and procedural knowledge and the relationship between them. The relationships between conceptual and procedural knowledge become important when knowledge is used dynamically in problem solving or in performing some nontrivial task (Silver, 1986). "When concepts and procedures are not connected, people may have a good intuitive feel for mathematics but not able to solve problems, or they may generate answers but not understand what they are doing" (Even, 1990, p. 527).

Knowledge about mathematics includes the knowledge about the nature of mathematics which is more that the procedural and conceptual knowledge. It refers to a more general knowledge about a discipline which guides the construction and use of both procedural and conceptual knowledge (Even, 1990, p. 527). It includes ways, means and processes by which truths are established as well as the relative centrality of different ideas (Lampert, 1988; Shulman, 1986; Thompson, 1984).

"The research into teachers' knowledge has often led to the conclusion that teachers were not able to interpret statistical concepts in an appropriate way or that teachers hold a poor knowledge towards specific statistical topics" (Eichler & Zapata-Cardona, 2016). The knowledge needed to teach statistics is regarded to be different than the knowledge needed to teach mathematics (Groth, 2007). Therefore, mathematics teachers' subject matter knowledge needed to teach statistics is also believed to be different than the subject matter knowledge needed to teach mathematics. One reason may be that although statistics utilises a lot of mathematics but it also involves nonmathematical activities. Hence, necessitating a requirement for mathematics teachers to have the subject matter knowledge involved in the enhancement of statistical literacy.



(Gal, 2002, p. 4)

Figure 2.5. Gal's model of statistical literacy

Gal (2002) in his model of statistical literacy suggested that there are five knowledge elements and two dispositional elements that together enable statistically literate behavior (p. 4) as shown in Figure 2.4.

The knowledge component consists of five cognitive elements: literacy skills, statistical knowledge, mathematical knowledge, context knowledge and critical questions whereas the dispositional component consists of two elements: critical stance and belief and attitudes. These knowledge and dispositional components are included as mathematics teachers' subject matter knowledge involved in the enhancement of statistical literacy.

The five elements of the knowledge component form the base and jointly contribute to teacher's ability to comprehend, interpret, critically evaluate and if required to react to statistical messages. These statistical messages can be presented in various forms such as oral, written text, tables or graphical representations. All the five knowledge elements do not function separately; these elements do overlap and are dependent of one another under the major component (Gal, 2002).

For example, the understanding of statistical messages requires activation of specific literacy skills (Mosenthal & Kirsch, 1998). To understand statistical message in the form of text requires various texts processing skills such as to derive the meaning from the stimulus presented to the readers, the written portion can be long that demand complex text comprehension skills or can be sometimes presented in the form of table or graphs with only few words which require the readers to comprehend the surrounding text in order to place the statistical part in the suitable context. The above clearly indicates that statistical literacy and general literacy are intertwined (Gal, 2002, p. 7).

The statistical knowledge base involves the knowledge of statistical and probabilistic concepts and procedures used to comprehend and interpret statistical messages. Statistical knowledge base consists of five parts:

> Knowing why data are needed and how data can be produced
> Familiarity with basic terms and ideas related to descriptive statistics
> Familiarity with basic terms and ideas related to graphical and tabular displays
> Understanding the basic notions of probability
> Knowing how statistical conclusions or inferences are reached (Gal, 2002, p. 10)

The third knowledge element needed for teachers to be statistically literate is the mathematical knowledge base. The mathematical knowledge base creates the awareness of the mathematical procedures underlying the production of common statistical indicators such as mean or variance. Even though understanding mathematical derivations is of some importance in presenting the key ideas of some statistical concepts but it will be wise to limit these mathematical derivations because of the availability of computers to do the computations. However, numeracy skills at certain sufficient level are still needed so that correct interpretations of numbers in statistical reports can happen (Gal, 2002, pp. 13-14).

Context knowledge base gives life to data. Context motivates procedures. Data should be viewed as numbers with context. Basis to any result interpretation will be the source of meaning and all these happens when context is introduced. Statistical messages are produced by commercial, political or other agendas for their needs which may be absent from typical statistics classrooms. Critical evaluation of such messages is necessary because the sources may not necessarily report balanced or objective findings (Gal, 2002, p. 15).

Teachers must examine the reasonableness of statistical claims presented in the media, be concerned about the validity of such claims, seek credibility of the evidence underlying the conclusions, and reflect upon possible alternative interpretations of conclusions (Wanta, 1997). Teachers must exercise their critical skills followed by critical questions where necessary. Gal (1994) mentioned that adults [teachers] should bear in mind the list of “worry questions” regarding any statistical information presented or displayed.

The dispositional aspects of statistical literacy are used for three related but distinct concepts that are important for statistical literacy- critical stance, beliefs and attitudes. Statistically literate teachers should have questioning attitude towards statistical messages that are misleading or incomplete. The personal list of “worry questions” should be activated when they encounter such messages. However, teachers’ critical stance is influenced by their own supporting beliefs and attitudes (Gal, 2002, pp. 15-16).

Despite the centrality of statistical literacy in various life contexts, the nature of the knowledge or skills and dispositions have not received detailed discussion in the literature (Gal, 2002). However, when concerning teachers, there are many factors that affect the ability of teachers to teach statistical topics in such way that improves the statistical literacy of learners and the subject matter knowledge is one of the factors (North, Gal, & Zewotir, 2014).

Summary. Teachers’ subject matter knowledge has a major influence on their teaching and students’ learning of the subject matter. The importance of teachers’ subject matter knowledge was highlighted by many researchers (Ball & McDiarmid, 1989; Even, 1990; Grossman, Wilson, & Shulman, 1989).

However, Shulman (1986, p. 7) highlighted that the subject matter knowledge had absence of focus in the research on teaching and referred to as the “missing paradigm” which might have serious consequences both for policy and for research.

Several studies have taken the call to explore further on teachers' subject matter knowledge. These studies explored on the domain of teachers' subject matter knowledge and suggested the components within the domain (Ball, Hill, & Bass, 2005; Ball, Lubienski, & Mewborn, 2001; Ball, Thames, & Phelps, 2008; Hill, Ball, & Schilling, 2008; Hill, Schilling, & Ball 2004).

Studies have also suggested that research on teachers' subject matter knowledge in mathematics should focus on specific topic. Even (1990) argued that, "analysing what teachers' subject matter knowledge means in general in mathematics, does not inform us of what subject matter knowledge teachers need to have in order to teach a specific piece of mathematics" (p. 522). Therefore, the researcher was interested in mathematics teachers' topic specific subject matter knowledge.

On the other hand, the subject matter knowledge needed to teach statistics is regarded to be different than mathematics. One reason is that statistics incorporates a lot of non-mathematical activities involving real life situations. Therefore, mathematics teachers should have the subject matter knowledge involved in real life situations if they are to meet the goal of preparing students to use statistics in their daily life. This further prompted the researcher's interest to focus on the subject matter knowledge of a specific statistics topic related to real-world situations. Consequently, the idea of the development of statistical literacy which emphasized a lot on real life situations drew the interest of the researcher.

Gal (2002) a prominent researcher in the field of statistical literacy mentioned that despite the centrality of statistical literacy in various life contexts, the nature of the knowledge and dispositions involved in statistical literacy did not receive much attention. He proposed through his model of statistical literacy that there are five knowledge and two dispositional elements involved in the development of statistical literacy. These elements also underpin the subject matter knowledge needed to teach statistics. Moreover, studies have highlighted that teachers' subject matter knowledge is one of the factors that affect the ability of teachers to teach statistical topics in such way that improves the statistical literacy of learners (North, Gal, & Zewotir, 2014).

Then again, the subject matter knowledge of mathematics teachers who are already involved in teaching and prospective mathematics teachers also exhibit differences. Mathematics teachers build upon their knowledge of subjects when they prepare to teach. Their subject matter knowledge undergoes transformation while they are involved in teaching. Their initial knowledge of content is further enhanced by other knowledge components such as knowledge of learners, curriculum, and teaching contexts (Grossman, Wilson, & Shulman, 1989).

Although mathematics teachers learn more about the subject matter when they come into teaching, they also require adequate background about their subject prior to their induction to teaching. Therefore, even as pre-service mathematics teachers, these teachers need to have a solid subject matter knowledge foundation upon which greater subject matter competence can be build (Grossman, Wilson, & Shulman, 1989). However, the question is of what this foundation should consist especially in relation to the enhancement of statistical literacy.

Hence, the researcher focused the present study on pre-service mathematics teachers' subject matter knowledge of a specific statistics topic involved in the enhancement of statistical literacy to gain better understanding of the subject matter knowledge foundation.

Research Related to Subject Matter Knowledge

Earlier research on teachers' subject matter knowledge have been general and not topic specific (Even, 1990). One study that focused on the topic specific subject matter knowledge was done by Even (1990). Even studied prospective secondary teachers' knowledge and understanding of functions. Her study found that teachers' knowledge of functions tends to be weak. She argued that simply assuming that teachers have a comprehensive and well-articulated knowledge of mathematics they have to teach is not right (Ball, 1988, 1990; Even, 1990). Instead, efforts must be made to improve mathematics courses needed for teachers to teach.

Teacher education should refer to topics included in high-school curriculum. Furthermore, teachers' subject matter knowledge with respect to the two aspects; "knowing that" and "knowing how" should be sufficiently comprehensive and well-articulated for teaching (Even & Tirosh, 1995).

Baturo and Nason (1996) found that first-year teacher education students' subject matter knowledge of area measurement was rather impoverished in nature. They seemed to hold on to assumptions that mathematics is mainly an arbitrary collection of facts and rules. Their mathematical knowledge is computational in nature and their mathematical ideas have little or no relationship to real objects. The study focused on four aspects of the subject matter knowledge; substantive knowledge, knowledge about the nature and discourse of mathematics, knowledge about mathematics in culture and society and dispositions towards mathematics.

In the area of statistics, research was done to describe the subject matter knowledge required to teach statistical association at the secondary level (Casey, 2008). The subject matter knowledge in Casey's study was expanded to include the triad; mathematical, statistical and context knowledge, the knowledge components of

Gal's (2002) model of statistical literacy. A practice based qualitative approach was taken in this study which is regarded unique because it was the first study that used practice-based approach to describe the subject matter knowledge needed for teaching statistics at the secondary level.

The findings of Casey's (2008) study showed that teachers need a substantial knowledge base for teaching the specific topic of correlation coefficient. It was found that it is necessary for teachers to know the "how" and "why" of concepts to meet the demands of teaching. The linking between the "how" and "why" of topics may be something that sets apart specialized mathematical knowledge for teaching statistical association from common mathematical knowledge of statistical association theorized by Hill, Schilling and Ball (2004), for teachers of statistics.

Summary. Previous studies related to topic specific subject matter knowledge in mathematics was found to be limited (e.g. Even, 1990; Baturo & Nason, 1996). Furthermore, research related to topic specific subject matter knowledge in statistics was also found to be very lacking. The researcher could identify only one study that looked at teachers' subject matter knowledge in statistics which focused on the topic of statistical association (Casey, 2008) and was aimed at in-service teachers.

Thus, there existed a gap in studies related to pre-service mathematics teachers' topic specific subject matter knowledge in statistics. This further encouraged the researcher to focus at pre-service mathematics teachers' subject matter knowledge of a specific statistics topic involved in the enhancement of statistical literacy to gain better understanding of the subject matter knowledge foundation.

Levels of Subject Matter Knowledge

Watson (1997) conducted a study on students' understanding of statistics associated to the development of statistical literacy and identified a framework of three tiers to assist in the development of statistical literacy.

The three-tiered framework originated from response to items based on media reports used in assessing statistical literacy. The first tier focuses on terminology which is interpreted broadly with the inclusion of graphical representations as well as specific phrases. Basic understanding of terminology is regarded as necessary before the context is involved (Watson, 1997, p. 2).

Context is the second tier of the framework. Context is the central idea that focuses on the understanding and also the application of both statistical and probabilistic language when it is embedded in a context of wider social discussion. Context offers deeper understanding of terminology. However, it is not necessary for complete understanding of terminology to take place before the context is introduced (Watson, 1997, p. 3).

The third tier of the statistical literacy framework focuses on the critical thinking aspects that enable people to challenge statistical statements that are made without proper statistical justification. However, to do this, the understanding of both terminology and the context associated with the earlier tiers is necessary. The ultimate goal of any statistical literacy program is to have learners reach the third level of the framework (Watson, 1997, p. 4).

Watson and Callingham (2003) further developed the three-tiered framework into a hierarchical construct to analyse students' levels of statistical understanding associated with statistical literacy. The sample was taken from five Catholic schools in the Australian state of Tasmania. A total of 673 students in Grades 5 to 10 were involved. Survey was administered by the classroom teachers during class time following instructions provided in an Administration Manual. The responses were coded using a code book. The coding schemes were hierarchical based on the structure of the response (Biggs & Collis, 1982, 1991), and the statistical appropriateness of the response (Watson, 1997). Six levels of understanding were categorized as follow:

Table 2.1

Hierarchical levels of statistical literacy

Level 1 (Idiosyncratic)	Idiosyncratic engagement with context, tautological use of terminology, and basic mathematical skills associated with one-to-one counting and reading cell values in tables.
Level 2 (Informal)	Only colloquial or informal engagement with context often reflecting intuitive non-statistical beliefs, single elements of complex terminology and settings, and basic one-step straightforward table, graph, and chance calculations.
Level (Inconsistent)	Selective engagement with context, often in supportive formats, appropriate recognition of conclusions but without justification, and qualitative rather than quantitative use of statistical ideas.
Level 4 (Consistent Non-critical)	Appropriate but non-critical engagement with context, multiple aspects of terminology usage, appreciation of variation in chance settings only, and statistical skills associated with the mean, simple probabilities, and graph characteristics.
Level 5 (Critical)	Critical, questioning engagement in familiar and unfamiliar contexts that do not involve proportional reasoning, but which do involve appropriate use of terminology, qualitative interpretation of chance, and appreciation of variation.
Level 6 (Critical Mathematical)	Critical, questioning engagement with context, using proportional reasoning particularly in media or chance contexts, showing appreciation of the need for uncertainty in making predictions, and interpreting subtle aspects of language.

(Watson & Callingham, 2003, p.118)

In 2006, a framework for analysing students' levels of understanding of linear measurement was developed by Barrett, Clements, Klanderman, Pennisi, and Polaki. They based the students' understanding on coordination of geometric reasoning and measuring strategies on a fixed perimeter task. The categories for students' levels of understanding of linear measurement:

Table 2.2

Levels of understanding linear measurements

Level 1	Visual guessing to assign length (naive unit strategy);
Level 2a	Inconsistent ways of identifying or iterating units: uses salient markers as a counting set for measuring;
Level 2b	Consistent identification or iterating of units;
Level 3a	Coordinating iterated-unit items, side-lengths, and collections of side lengths to obtain perimeter;
Level 3b	Coordinating length attributes, yet with further tendency and ability to relate multiple cases.

(p. 197)

The sample consisted of 38 students ranging from grade two to grade ten. The students were asked to draw and label all the rectangles (triangles) with a fixed perimeter of 24 units and explain their answers. The study found that four, eight, and two from Grades two to three students exhibited levels 1, 2a and 2b respectively. One, five, five, and two Grade 5 to 6 students exhibited levels 1, 2a, 2b, and 3a respectively. One, seven, and three Grade 8 to 10 students exhibited levels of 2b, 3a, and 3b respectively.

Hill, Ball, Bass, and Schilling developed a coding rubric for measuring the mathematical quality of instruction through Learning Mathematics for Teaching (2006) project. The coding rubric consists of 83 codes grouped into five sections as follows:

> Section I: Instructional format and content, Section II: Knowledge of mathematical terrain of enacted lesson, Section III: Use of mathematics with students, Section IV: Mathematical features of the curriculum and teacher's guide, and Section V: Use of mathematics to teach equitably. (p. 6)

"Section II codes the teacher's knowledge of the mathematics entailed in the lesson as revealed by its enactment" (p. 7). This section consists of 12 codes which are grouped into five subsections as follow: "(a) teacher's use of language (3 codes), (b) examples and models used to represent mathematical concepts (4 codes), (c) degrees of mathematical explanation (3 codes), (d) development of mathematical elements (1 code), and (e) computational errors or other mathematical oversight (1 code)" (p. 7). Finally, "there is a global code used to record the coders' impression on teacher's level (low, medium, high) of mathematical knowledge. Overall, this section is designed to capture the teacher's understanding of the content being taught and the mathematical resources used during the lesson" (p. 7).

Each mathematical element in the coding rubric is coded as "present and appropriate (PA)", "present and inappropriate (PI)", "not present and appropriate (NPA)", or "not present and inappropriate (NPI)". Random pairs of researchers were assigned to code each videotaped lesson during The Learning Mathematics for Teaching project. Each lesson was coded individually by the coders and then followed by an overall level of the teacher's knowledge of mathematics as low, medium, or high. The level of teacher's knowledge was based on the coders' impression of the teacher's level of mathematical knowledge. Then, the coders met and reconciled their codes before giving their final level of mathematical knowledge.

Summary. This section included a discussion on the studies related to levels of subject matter knowledge. Watson (1997) developed a three-tiered framework for analysing statistical understanding in relation to the development of statistical literacy which later Watson and Callingham (2003) developed the framework into a six-level hierarchical construct. The framework and hierarchical construct were used to analyse statistical understanding and the development of these understandings in relation to statistical literacy. Similarly, Barrett, Clements, Klanderman, Pennisi, and Polaki (2006) developed a framework for analysing students' levels of understanding of linear measurement.

Since the present study was aimed at investigating the presence or existence of the subject matter knowledge through what constitutes such knowledge, therefore the researcher decided that the pre-service mathematics teachers' level of subject matter knowledge of measures of central tendency in the present study would be determined using the coding rubrics adapted from The Learning Mathematics for Teaching (LMT) (2006) project developed by Hill, Ball, Bass, and Schilling. The coding rubrics developed from this project determined the level of knowledge by taking into account the existence of such knowledge elements.

Research Related to Malaysian Pre-service Teachers' Subject Matter Knowledge

Several studies have explored Malaysian pre-service teachers' subject matter knowledge. Nik Azis (1995) reported that the pre-service teachers in his study focused a lot on computational skills that they were left with no time to guide their students to construct meaningful basic concepts in mathematics. The pre-service teachers believed that the most suitable approach to teach mathematics is through computational approach because they viewed mathematics as a collection of rules that is used to find correct answers for the given questions.

According to Nik Azis (1995), students were not encouraged to use their own strategies to solve mathematical problems and were treated as passive receivers of mathematics. Students were trained to memorize certain mathematical facts and trained to master certain mathematical skills. Demonstration-practice method was used by the pre-service teachers where they demonstrated how to solve certain questions followed by the students asked to solve certain questions from either the textbook or workbook based on the shown examples. The demonstration-practice cycle is repeated if the pre-service teachers find students still are unable to solve the assigned questions.

Nik Azis (1995) also revealed that these pre-service teachers used the school textbook and past year examinations book as their only resources for teaching school mathematics. However, the pre-service teachers had issues related to effective use of textbook and they followed the sequence in the textbook strictly without attempting modification. These pre-service teachers also exhibited lack of confidence towards certain school mathematics topics.

On the other hand, trainee teachers in the teacher training institutes (formerly known as teacher training colleges) had demonstrated a poor understanding of mathematical concepts and a lack of mathematical skills (Cheah, 2001; Koe, 1992; Ng, 1995). Cheah (2001) used a multiple case study to explore the mathematical beliefs of six trainee teachers who majored in either mathematics or science. The findings revealed that these trainee teachers' central belief about mathematics was already shaped during their schooling experiences. The trainee teachers held on the belief that mathematics was procedural and that learning of this subject was focused on procedures, algorithms, and the use of formulae. Koe (1992) revealed that trainee teachers had difficulty answering mathematics questions taken from primary six standardized examinations exhibiting a lack of mathematics content. Similar finding was reported by Ng (1995).

Nik Azis (2003) emphasized that mathematics teachers must know and understand the school mathematics that they are teaching. Nik Azis pointed out that:

> Currently, there is a tacit assumption that, by the time pre-service secondary mathematics teachers have completed their university mathematics courses, they will have the understanding of school mathematics subject matter required for teaching the subject matter effectively. However, recent research has shown that this assumption might not be valid. We cannot take for granted that teachers' knowledge of the content of school mathematics is in place by the time they complete their secondary school learning experiences. (p. 6)

Therefore, mathematics teachers must be given the opportunity to revisit school mathematics in such way that allows them to develop deeper understandings of mathematics. However, Nik Azis (2003) mentioned that in order to develop these understandings is not easy:

> It will be challenging to help teachers understand school mathematics content at a deeper conceptual level, to help them understand the big ideas of mathematics and to be able to present mathematics as a unified discipline, a woven fabric rather a patchwork of discrete topics, and help them developing mathematical reasoning. At the moment, this kind of knowledge is beyond what most mathematics teachers experience in pre-service mathematics courses. (pp. 7-8)

Wun (2010) attempted to study pre-service mathematics teachers' subject matter knowledge in perimeter and area. A multiple case study was used and data was collected using clinical interview. Data was obtained from eight pre-service secondary school mathematics teachers enrolled in a Mathematics Teaching Methods course at a public university. Five basic types of knowledge of perimeter and area were investigated namely conceptual knowledge (CK), procedural knowledge (PK), linguistic knowledge (LK), strategic knowledge (SK), and ethical knowledge (EK).

The findings of the study revealed that all the pre-service teachers lacked in the conceptual knowledge underpinning the formula for the area of a rectangle. However, in regards of the procedural knowledge most of them held a better procedural knowledge of calculating perimeter and area of composite figures.

Wun (2010) also revealed that in terms of linguistic knowledge (LK), the pre-service teachers used appropriate mathematical symbols to write the formula for the area, used appropriate mathematical terms to justify their selection of shapes that have a perimeter and an area, and understand general measurement convention. However, these pre-service teachers showed limited knowledge about the conventions related to writing and reading of Standard International (SI) area measurement units. Most pre-service teachers did not attempt to check the correctness of the answers for the perimeters and areas.

Wun (2010) also found that only one pre-service secondary school mathematics teacher secured a high level of subject matter knowledge which leads to a conclusion that these pre-service teachers lack subject matter knowledge of perimeter and area that they are expected to teach.

Effective mathematics teachers must have a good command of both school mathematics and university mathematics, adequate effective pedagogical practices preparation, and achieve high overall academic performance (Nik Azis, 2003). This is found crucial because the success of students depends very much on the quality of teachers. Subject matter knowledge and mathematical discourse are important aspects in the mathematics teacher preparation (Nik Azis, 2008).

Summary. In Malaysia, studies have found that pre-service mathematics teachers had issues related to the subject matter knowledge such as poor understanding of the mathematical concepts, lack of mathematical skills, difficulty in answering mathematics questions, and weak in mathematics content. However, in the Malaysian context not many studies have actually looked at topic specific subject matter knowledge. Since there existed a lack in studies which investigated pre-service mathematics teachers' topic specific subject matter knowledge especially in statistics, therefore the present study attempted to fill this gap.

Research Related to Students' Understanding of Measures of Central Tendency

Initial studies on students' understanding of averages were focused on tertiary students. Pollatsek, Lima and Well (1981) conducted a study on tertiary students who attempted to solve problems involving the appropriate weighting and combining of means into overall mean. The study found that when tertiary students were given means for two unequal-sized samples and asked to find the mean of the combined sample, the students tended to weight the samples equally and find the midpoint of the two-sample means. This is referred as “closure misconception” which students already have in the back of their mind when they operate with means and variance in statistics (Mevarech, 1983). The results also indicate that most of the students do not understand the concept of the weighted mean. Students dealt with the mean in a rather computational act than a conceptual act.

Another similar study on university students was done by Mevarech (1983). The study found that one reason that contributes to the students’ difficulties in understanding both simple and weighted average problem is that they think of the operations of the average as binary operations satisfying the four laws of an additive group.

Pollatsek, Lima and Well (1981) suggested that three additional kinds of knowledge exist in the schema of the mean: (1) functional, (2) computational, and (3) analog knowledge. Pollatsek et al. referred *functional knowledge* to the understanding of the mean as a meaningful real-world concept. When the mean has a real-world referent, additional knowledge is needed in determining the choice of numbers to be entered into a computational formula. *Analog knowledge* involves visual or kinesthetic images of the mean as a ‘middle’ or balance point.

Some of the later studies attempted to study younger children’s understanding of average. One of the first studies reported on younger children was by Strauss and Bichler (1988). Strauss and Bichler (1988) conducted a study to determine the development of children’s understanding of seven properties of the average (arithmetic mean):

a. The average is located between the extreme values;
b. The sum of deviations from the average is zero;
c. The average is influenced by values other than the average;
d. The average does not necessary equal one of the values that was summed;
e. The average can be a fraction that has no counterpart in physical reality;

f. When one calculates the average, a value of zero, if it appears, must be taken into account;
g. The average value is a representative of the values that were averaged.
(Strauss & Bichler, 1988, p. 66)

The sample consisted of twenty children of the ages of 8, 10, 12, and 14 years. Students in the study were asked a series of structured questions in a one-on-one setting to probe students' understanding on the properties of average (arithmetic mean). Their understandings were probed on "statistical" and "abstract" properties of average as well as "average as a representative of a set of values".

The analysis of the study identified two very different levels of difficulty in their tasks. The study found that students were aware that the mean was between the extremes and that particular data values can influence the mean. However, students had extreme difficulty in measurement ideas such as minimizing deviations or a zero-data value must also be included and accounted in mean calculation.

Strauss and Bichler (1988) suggested that the reasoning underlying children's understanding of the properties in question should be accounted in educational practice because that produces the behaviour we see when children solve problems. The authors also concluded that the understanding of students increases with age. This presents evidence to support the existence of levels of development for different properties of the concept of average.

In another study, Mokros and Russell (1995) attempted to understand the characteristics of fourth through eight graders' constructions of "average" as a representative number summarising a data set. Twenty-one students were interviewed using a series of open-ended problems. The open-ended problems went beyond straightforward algorithmic computations and allowed children to construct their own notion of representativeness. The type of tasks Mokros and Russell used required students to work backwards from a mean to possibilities for a data set that could have that mean.

The study identified five approaches used by fourth, sixth and eighth graders in considering the meaning of average: average as mode, average as algorithm, average as reasonable, average as midpoint, and average as mathematical balance point. The results of the study revealed that students focused on *modes* in data sets had difficulty in working backwards from the mean to construct a data set if they were not allowed to use the actual average value itself as a data value.

These findings concluded that the modal thinking students do not see the whole data set and its distribution as an entity itself. This finding is found to be similar to Cai (1995) whereby students in this study had no problem calculating a mean when the data is provided but they had difficulty working backwards when they were asked to fill the missing values of the data for the given mean. Mokros and Russell also discovered that students who held algorithmic conception of average were unable to make connections from their computational procedures back to the actual context. The premature introduction of the algorithm may have caused "short circuit" in children's reasoning about the average.

Cortina, Saldanha and Thompson (1999) elaborated on a multiplicative conception of the arithmetic mean that is grounded in quantitative reasoning. They investigated middle-school students' conceptions as these students engaged in problem tasks that were intended to support their understanding of the arithmetic mean multiplicatively. The findings of the study showed that the students demonstrated sophisticated multiplicative reasoning. These students' understanding of "average" was more than procedural. The "fair share" conception by Mokros and Russell (1995) was also found to be consistent. The results of the study suggested that students should conceive first the idea of group performance as a measurable attribute before the idea of the mean as an appropriate measure of that attribute is brought in.

Watson and Moritz (1999, 2000) conducted a longitudinal study on the development of the understanding of average in students ranging from grades 3 to 9. The study probed students' understandings of average in media or every day contexts. Students were also asked to work backwards where they were asked to fill in missing data values when given the mean and to find the average in weighted mean situations.

SOLO taxonomy was used to model student responses on tasks about average. Six levels of student understanding of average were described: *pre-structural, uni-structural, multi-structural, relational, application of average in one complex task* and *application of average in two complex tasks.* The study provided strong evidence that students' conceptions of average followed a developmental path.

The follow-up interviews that were conducted 3 to 4 years later with the same students also found that these students maintained their level of thinking about average. It is believed that once students had developed powerful and flexible conceptions of the mean which included weighted averages in real contexts, the students maintained the conceptions over time. However, to develop students' concept of the mean to the point that it is a representative of a data set might take many along time (Watson & Moritz, 2000).

The study found evidence that supports students' development of understanding of the average highlights the importance of teachers' possessing conceptual understanding of the average. It is believed that teachers who attained good conceptual understanding of average were able to assist students to progress from pre-structural levels to higher levels.

Konold and Pollatsek (2002) conducted a study on the complexity of the average concept. Their study extends Mokros and Russell's (1995) work. Four conceptual perspectives for the mean were postulated: mean as *typical value*, mean as *fair share*, mean as *data reducer*, and mean as *signal amid noise*. However, in their study they argued that out of the four conceptual perspectives, the perspective that the mean as "signal amid noise" is regarded as the most important and useful conception because they felt that this conception of the mean is necessary in comparing two data sets.

The study suggested that the context of comparing data sets using mean should be introduced to students and also suggested that the conceptions of the mean as typical or fair share should not be emphasized to students because of these conceptions incapability in group comparisons.

Konold and Pollatsek's conceptions of the mean as typical and fair share are closely connected to data analysis in statistics, whereas the mean as signal and data reducer are closely connected to decision-making in statistics. However, Konold and Pollatsek's conceptions of the mean as signal or data reducer might be useful in a statistician's perspective but looking at the perspective on students' primary intuitions of the mean, mean as typical and mean as fair share are regarded important introductions to the notion of measures of central tendency (Mokros & Russell, 1995; Watson & Moritz, 1999, 2000).

McGatha, Cobb and McClain (2002) reported on their analysis of performance assessment tasks administered in a seventh-grade classroom. The assessments were aimed to obtain data on students' current statistical understandings. The tasks were designed to provide information about students' current understandings of creating data, organizing data, and assessing the center and "spreadoutness" of data.

The study found that regardless of the task situation, students naturally viewed the mean as a procedure which is used to summarise a set of numbers. These students held on a perspective that data analysis is "doing something with the numbers." The study suggested that classroom teaching should support a shift in students' reasoning towards data analysis as inquiry rather than procedure

Batanero, Cobo, and Diaz (2003) assessed secondary school students' personal meaning they attribute to mean, median and mode using a questionnaire that made up of 9 open-ended tasks. The open-ended tasks provided detailed reasoning to student responses.

The study suggests that understanding a concept is a continuous constructive process. Students progressively acquire and relate the different elements of the meaning of the concept. This is believed to emerge from students' meaningful practices linked to repeated solution of problems that are specific to that concept. Students who are exposed to these repeated activities of solving significant problems related to the concept will progressively acquire and widen their understanding of the concept.

Garcia and Garret (2006) attempted to analyse how students act when faced with open-ended questions that are closely related to multiple-choice questions. They wanted to find out if those students who chose the correct options in multiple-choice questions have done so using a clear criterion. This was based on their observations on the students' actions during the process of solving open-ended questions. However, the results of the study found that many students who chose the correct answer for the multiple-choice questions are not able to demonstrate reasonable methods for solving open-ended questions.

In another study, Sirnik and Kmetič (2010) attempted to find out Slovenian students' understanding of the concept of arithmetic mean. The study is based on a sample of two different age groups of learners and focused on knowledge of the concept and computational rule for arithmetic mean, distinguishing among means, and how the concept of average is understood in the context of a given problem.

The findings of the study revealed that although students master computational algorithms for statistical means but they were still unable to apply them to explain everyday situations. Students also lacked experience in open-ended tasks in mathematical and everyday contexts, statistical investigations, reverse questions in mathematical and everyday contexts, and questions fostering reasoning and decision making, including explanation.

Watson and Chick (2012) analysed the responses of 247 middle school students to items requiring the concept of average in three different contexts. The outcome of the analysis of these three contexts has resulted in two implications for teaching. One explicit engagement with the concepts is needed suggesting that teachers need to be able to devise remedial strategies to assist students to come to a stronger understanding. In addition, the limited nature of some students' responses suggests that teachers should encourage in-class discussions on how statistical information complements and develops understanding of real-world contexts.

Summary. Studies regarding students' understanding of average covered a wide range of aspects. Several initial studies focused on the weighted mean problems that involved tertiary students (Pollatsek, Lima, & Well, 1981; Mevarech, 1983).

Later, the focus shifted to younger children that studied their development in the understanding of the properties of the average (Strauss & Bichler, 1988).Mokros and Russell (1995) attempted to study children's constructions of average as a representative number summarising a data set. Similarly, Cortina, Saldanha and Thompson (1999) elaborated on the multiplicative conception of the arithmetic mean. Later Konold and Pollatsek (2002) extended the work of Mokros and Russell (1995) and further attempted to study the complexity of the average concept. These studies in general indicate that students held better procedural than conceptual understanding of the average.

On the other hand, some later studies highlighted the role of teachers in the development of students' understanding of average. For example, Watson and Moritz's (1999, 2000) longitudinal study on the development of the understanding of average in everyday contexts highlighted the importance of teachers' possessing good conceptual understanding of average if they were to assist students in their attainment of the conceptual understanding of average.

Meanwhile, Batanero, Cobo, and Diaz (2003) assessed students reasoning about mean, median, and mode; suggested that meaningful practices are linked to students understanding of the concept. Another recent study by Watson and Chick (2012) also highlighted the importance of teachers in developing students' understanding of the average in meaningful real-world contexts.

Since several studies have mentioned the importance of teachers in students' understanding of the average, therefore the next section included a discussion on studies related to teachers' understanding of measures of central tendency.

Research Related to Teachers' Understanding of Measures of Central Tendency

Measures of central tendency have been part of the mathematics curriculum even before the introduction of statistics in the curriculum. One reason may be that measures of central tendency appear quite frequently and became part of our daily life.

Research pertaining to teachers' understanding of the measures of central tendency is found to be limited (Jacobbe & Carvalho, 2011).

Among the studies that focused on teachers' understanding of the measures of central tendency are some that focused mainly on the arithmetic mean (e.g. Batanero, Godino, & Navas, 1997; Gfeller, Niess, & Lederman, 1999).

Batanero, Godino, and Navas (1997) analysis of the responses to a written questionnaire conducted on 367 pre-service teachers in Spain revealed the existence of conceptual errors about elementary statistical concepts. The results revealed that there was a lacking on teachers' understanding of the algorithm for calculating the arithmetic mean; teachers were also ignorant of the relationship between mean, median, and mode; they were unable to discern properly that a value which is a typical for a given context and they also showed confusion in determining the positions of mean, median, and mode in asymmetrical distributions. The pre-service teachers also exhibited little or no understanding of the effect of outliers on mean.

Gfeller, Niess, and Lederman (1999) conducted a study on teachers' views and the representations that they use in solving problems involving arithmetic mean in United States of America. 13 mathematics and 6 science pre-service secondary school teachers responded to an instrument which consisted of seven concrete and three abstract questions on arithmetic mean. Their responses were classified according to whether the participants utilised a computational algorithm, a procedure-based understanding, or knowledge of variation in solving the problems related to the mean.

The results revealed that there was no significant difference in the use of computational algorithm between mathematics and science pre-service teachers but there was a significant difference regarding the use of variation to create a balance point with mathematics pre-service teachers' keener towards the utilisation of this approach as compared to science pre-service teachers. In fact, all the mathematics pre-service teachers utilised this approach. In the context of the teachers' views of the mean, majority of them held multiple views.

The results of Gfeller, Niess, and Lederman's (1999) study revealed the opposite compared to Batanero, Godino, and Navas's (1997) study where the participants Gfeller et al.'s study was found to hold a deeper conceptual understanding of the mean as compared to the participants of Batanero et al.

Hobden (2014) conducted a study that focused on the concept of median given in daily life context. Her study was to determine the levels of statistical literacy evident in 316 prospective teachers' explanations of the median in the context of HIV/AIDS survival times. Hobden used a categorization framework drawn from Watson's (1998, 2006) three-tiered statistical literacy hierarchy and the SOLO taxonomy (Biggs & Collis, 1982). The findings of this study revealed that more than half of the teachers were classified below the basic understanding of the median and only about two percent of the participants were classified had coherent explanation of the median in the given context.

Several studies focused on teachers' understanding of all three measures of central tendency (e.g. Groth & Bergner, 2006; Jacobbe, 2007, 2008).

Groth and Bergner (2006) conducted a study on 46 pre-service elementary and middle school teachers. The pre-service teachers responded to a written item pertaining to the measures of centre that asked the pre-service teachers to compare and contrast the statistical concepts of mean, median, and mode. The data produced was examined using the Structure of the Observed Learning Outcome (SOLO) taxonomy (Biggs & Collis, 1982, 1991) and Ma's (1999) conception of Profound Understanding of Fundamental Mathematics (PUFM).

PUFM is described as "an understanding of the terrain of the fundamental mathematics that is deep, broad, and thorough" (Ma, 1999, p. 120). PUFM exhibits the connected knowledge of both concepts and procedures. Ma (1999) noted that there are four different ways teachers can attain PUFM: studying teaching materials intensely, learning mathematics from colleagues, learning mathematics from students, and learning mathematics from doing it. All these happened within the context of teaching environment which may be due to the fact that the need is more apparent in the context of teaching rather than in the context of taking a course at the university.

Groth and Bergner (2006) revealed that there existed four distinct levels of thinking in view to pre-service teachers' abilities to compare and contrast the concepts of mean, median, and mode: unistructural/concrete symbolic, multistructural/concrete symbolic, relational/concrete symbolic, and extended abstract. The findings revealed that most of the pre-service teachers exhibited thinking which was very much similar to the thinking of elementary and middle school teachers as reported from the literature (Cai, 2000; Mokros & Russell, 1995; Watson & Moritz, 2000).

8 pre-service teachers fell in the unistructural/concrete symbolic level as their responses only involved the recitation of definitions in describing the measures of centre. 21 teachers exhibited the multistructural/concrete symbolic level of thinking because their responses presented the view that the calculations or findings of measures of centre provided some type of value but they were not sure what value the calculations/findings actually provided. Thirteen teachers fell in the relational/concrete symbolic level of thinking because their responses were beyond calculation of the measures which included discussion of what is "typical" about a certain set of data. Finally, only 3 teachers exhibited more abstract levels of thinking in which their

responses went beyond the procedural knowledge and included a discussion on the appropriateness of one measure of centre in describing a certain set of data.

The study suggested that since learning within the context of teaching is another important aspect of developing PUFM, it would be worthwhile to investigate approaches for integrating content-knowledge learning within the context of teaching in teacher preparation programs.

Meanwhile, Jacobbe (2007, 2008) conducted a case study on three elementary school teachers' understanding of average. Participants of this study were presented with three different distributions (one skewed to the left, one skewed to the right, and one normal). They were asked to arrange the distribution in ascending order for the values mean, median, and mode and were specifically asked to provide an example of when one measure of centre would be more representative for a given set of data than the other. The findings of this study revealed that even though some teachers had difficulty in the applying algorithm to different contexts but they were able to use the shape of a distribution in deciding when one data set would have greater mean, median, and mode compared to the other.

Santos and da Ponte (2013) conducted a study to understand the meanings that prospective elementary and kindergarten teachers gave to central tendency measures during statistical investigation. Both innovative and context-based interpretations were observed which showed the real understanding of the central tendency concepts. The findings of the study showed that these teachers had a large variety of interpretations of central tendency measures in which some were appropriate and the others not. For instance, the appropriate interpretations given to the concept of the mode are value associated to the highest frequency, "the majority…or most", and value referred or appeared more times whereas the "highest value" was found to be a problematic interpretation.

For the mean, "value that represents equilibrium or balance point" and fair share model meanwhile for the median, "value that divides the ordered data" and "up to the median there is 50% of the sample…" were found to be appropriate interpretations for the mentioned measures. The study also reported that their teachers showed confusion about the meaning of the measures and could distinguish among different measures such as there were problematic interpretations related to the confusion between the mean and the median with spread measures.

Another important term that is always associated with the measures of centre is "average". Average is commonly used in our daily life and understanding the association of this term with measures of centre is crucial. Several studies had focused on the general concept of average (e.g. Russell & Mokros, 1991; Callingham, 1997; Begg & Edwards, 1999; Estrada, Batanero, & Fortuny, 2004; Leavy & O'Loughlin, 2006).

Russell and Mokros (1991) conducted a study on students which also included 8 teachers who were teaching grade 4 to grade 8 and two mathematics coordinators. The study compared teachers' understanding of average to students and found that teachers exhibited the following ideas: average as modal, average as midpoint, and average as an algorithm relationship with majority of the teachers fell in the category of "average as an algorithm relationship". Average as what's reasonable was not exhibited by the teachers in this study. The teachers also showed varying levels of success with their application of the algorithm in finding the mean value. Russell and Mokros revealed that "introduction of the algorithm as a procedure disconnected from students' informal understanding of mode, middle, and representativeness which causes short-circuit in the reasoning of many children [adults]" (1991, p. 313).

Callingham (1997) investigated teachers' multimodal functioning in relation to the concept of average emphasizing on the arithmetic mean. A four-question survey was employed on 100 pre-service and 36 in-service teachers. The first question is on identifying the average given a set of data meanwhile the second and third questions involved participants using a graphical display of data to determine if one group performed better than the other group. The fourth question is on the concept of weighted mean. The SOLO Taxonomy with multimodal functioning developed by Biggs and Collis (1991) was the theoretical model used as the basis of the analysis.

The study revealed that the question on weighted mean was regarded difficult as compared to the first three questions. The findings of this study suggest that students at all levels need to experience concepts presented in a variety of contexts so that they could develop a range of strategies for problem solving and the skills to choose a suitable procedure. Therefore, the findings of this study suggest that teachers themselves need to experience similar types of activities in their pre- and in-service trainings. Teachers need opportunities to move from different modes of thinking and reflect on the processes they use in order for them to be able to develop a pedagogy that allows the similar to happen in the classroom.

In 1999, Begg and Edwards conducted a study on teachers in New Zealand. The study investigated 22 in-service and 12 pre-service elementary school teachers' ideas about teaching statistics that included some ideas related to average using unstructured, semi-structured and clinical interviews. The findings revealed that most teachers were not familiar with the mathematical definitions of mean, median, and mode. The common response for the word average "was in the middle". However, regarding their understanding of specific measures, the teachers possessed better understanding of the mean as compared to the median or the mode.

Estrada, Batanero, and Fortuny (2004) conducted a study on 367 pre-service primary teachers and attempted to assess their statistical knowledge. The instrument used in this study had selected items that specifically addressed on what the pre-service teachers would be required to teach. The study found that despite a majority of the teachers who had training in statistics but many still exhibited errors concerning the average. The common error was not being conscious of the effect of outliers on the mean. The teachers also carried out routine application of the algorithm in finding the mean value without taking into account the context and were not able to invert the algorithm. Teachers also exhibited confusion regarding mean, median, and mode. This may be due to the training received in the teacher preparation program which focused on calculation rather than the application of the concepts. The study recommended that future teachers should take statistical courses during their preparation programs that emphasize the application of the concepts.

Leavy and O'Loughlin (2006) conducted a study in Ireland on 263 pre-service elementary school teachers' understanding of the mean. The participants responded to five tasks and in order to further capture information on the participants' conceptual knowledge; clinical interviews were held with only 25 participants. The findings of the study revealed that the participants were confused with the concept of the mean with the other measures of centre. Some participants viewed the term "mean" as synonymous with "average".

Finally, one recent study by Koleza and Kontogianni (2013). This study was conducted on 166 first year pre-service teachers that investigated the level of statistical literacy of these pre-service teachers in their first-year university and after the end of schooling. The study adapted the framework of statistical literacy of Watson (1997, 2003) and Gal (2002). A modified SOLO taxonomy was used for the assessment of the participants' responses. The questionnaire items used covered several constructs which included the notion of average. The study reported that more than half of the participants attained the relational level with correct justification for the items on the notion of average.

On the other hand, it was found that not many studies have focused on teachers' professional knowledge for the measures of central tendency.

One study that focused on teachers' didactic knowledge was done by Cai and Gorowara (2002). The study investigated teachers' conceptions and constructions of pedagogical representations for teaching the concept of arithmetic average. The participants were 12 inexperienced and 11 experienced teachers. Data for this study was collected in three different ways: (1) teachers were asked to write an introductory lesson plan on average, (2) the teachers were asked to think how 6^{th} and 7^{th} grade students might solve each of the five problems related to arithmetic mean, and (3) the teachers were provided with six student responses for two of the five problems and were asked to evaluate the response using five-point scoring rubric. The data collected from three different sources contributed information for three different aspects of representations: (a) general pedagogical representations for classroom instruction, (b) knowing students' representations and strategies in problem solving, and (3) evaluating students' representations and solution strategies.

The results of this study revealed that the experienced teachers held better knowledge of student representations as well as common student errors compared to the inexperienced teachers. In the task requiring teachers to evaluate student solutions, both the experienced and inexperienced teachers showed similar results. Teachers gain experience and take on leadership roles are able to reflect upon their students' misconceptions to enhance their own understanding as well as their ability to address the misconceptions students reveal. However, this study also revealed that although the experienced teachers showed a more extensive and diverse knowledge of student representations but this was not clearly evident in their lesson plans. This scenario may indicate that some disconnectedness existed between teachers' knowledge and their planning.

Olfos and Estrella (2010) conducted a study on teachers understanding of teaching statistics. Their study addressed the relationship between the pedagogical knowledge of statistics in primary teachers in Chile and its impact on students learning. The study involved thirty-one 4th and 7th grade teachers from different schools of Valparaiso district. The teachers were tested on three questions, after reading a short text referred to non-typical values and properties of the mean. The teachers' answers provided evidence of their understanding of statistics and their pedagogical content knowledge.

The results of the study revealed that the teachers involved did not observe data within a context. Their procedural knowledge was not integrated with their conceptual knowledge about the mean and they had not been familiarized with the conceptual aspects of the mean. The involved teachers demonstrated rigid understanding of statistics, as part of mathematics, like a formal discipline. The study provided evidence that teachers who have limited content knowledge might not be able to follow the curricular guidance.

Another recent study on teachers' professional knowledge conducted by Watson and Callingham (2013) investigated teachers' pedagogical content knowledge about the concept average. The sample consisted of 26 teachers who were in their final year of a 3-year longitudinal project providing professional learning in statistics education from three Australian states. The teachers had teaching experience varied from two to more than 25 years and taught Year 5 to 12 with the main experience at the middle school level. The teachers completed teacher survey items on average.

Teachers in this study displayed no misunderstanding of the average but revealed that teacher-centred perspective taken in many of the responses that covered the content neglecting students as learners. The study suggested that teachers both at pre-service and in-service need help to develop pedagogy that introduces content in a way that students will understand and be successful in grasping the learning outcome.

Summary. Research pertaining to teachers' understanding of measures of central tendency was found to be limited (Jacobbe & Carvalho, 2011). Some studies have focused on teachers' understanding of arithmetic mean (e.g. Batanero, Godino, & Navas, 1997; Callingham, 1997; Gfeller, Niess, & Lederman, 1999; Leavy & O'Loughlin, 2006).

Whereas, several other studies have focused on the general concept of average (e.g. Russell & Mokros, 1991; Callingham, 1997; Begg & Edwards, 1999; Estrada, Batanero, & Fortuny, 2004; Leavy & O'Loughlin, 2006). However, limited studies have focused on all three measures of central tendency (e.g. Groth & Bergner, 2006; Jacobbe, 2007, 2008). Then again, only a handful of studies have focused on teachers' professional knowledge (e.g. Cai & Gorowara, 2002; Olfos & Estrella, 2010; Watson & Callingham, 2013).

Studies related to teachers' professional knowledge which focused on the pedagogical content knowledge revealed the existence of disconnectedness between teachers' knowledge and their teaching planning; they had limited content knowledge and; they needed help to develop pedagogy that introduces the content of measures of central tendency in ways that students will understand and grasp successfully the learning outcomes (Cai & Gorowara, 2002; Olfos & Estrella, 2010; Watson & Callingham, 2013).

These findings indicate an underlying issue related to teachers' subject matter knowledge of measures of central tendency because according to Shulman (1986) pedagogical content knowledge goes into the dimension of subject matter knowledge for teaching.

Furthermore, studies that focused on teachers' professional knowledge were mostly aimed at in-service teachers (e.g. Cai & Gorowara, 2002; Olfos & Estrella, 2010; Watson & Callingham, 2013). For instance, Jacobbe (2007, 2008) and Russell and Mokros (1991) studied in-service elementary and middle school teachers' understanding of averages. Whereas, Gfeller, Niess, & Lederman (1999) studied the views and representations used by teachers in problem-solving involving arithmetic mean.

Thus, there existed a gap in research related to pre-service mathematics teachers' subject matter knowledge of measures of central tendency and this prompted the researcher towards investigating pre-service mathematics teachers' subject matter knowledge of measures of central tendency. The researcher also had questions regarding the nature of subject matter knowledge of measures of central tendency held by Malaysian pre-service mathematics teachers.

Measures of Central Tendency in the Malaysian Mathematics Curriculum

In the Malaysian mathematics curriculum, students start to learn measures of central tendency concept formally in Year Five. Average is introduced in the Data Handling topic. Students learn to understand and use the vocabulary related to average. The algorithm to calculate average is also introduced at this level. Students also learn to use and apply the knowledge of average. Students are taught to calculate average for two or three numbers. Students are also introduced to real life problems involving average (Integrated Curriculum for Primary Schools Curriculum Specifications Mathematics Year Five, 2006).

In Year Six, students learn further to understand and compute average for more than three numbers. Students are further exposed to more problem- solving in real life contexts. Among the suggested teaching and learning activities are on problems involving average in simple text, tables or pictures. Polya's four step model is used in the problem-solving activities (Integrated Curriculum for Primary Schools Curriculum Specifications Mathematics Year Six, 2006).

In the Malaysian secondary mathematics curriculum, measures of central tendency are formally introduced in Form Three. However, in Form Two, students learn the fundamental idea of the measures of centre which is data. Students learn to understand and use the concepts of mean, median, and mode to solve problems which are confined to ungrouped data. Some of the suggested teaching and learning activities are on the use of sets of data from everyday situations, discussion on the suitable measurement in different situations and discussion on the appropriate use of the mean, median, and mode in certain situations (Integrated Curriculum for Secondary Schools Curriculum Specifications Mathematics Form Three, 2006).

Students continue to learn measures of central tendency in Form Four focusing on the grouped data. Real life situations are also emphasized at this level (Integrated Curriculum for Secondary Schools Curriculum Specifications Mathematics Form Four, 2006).

A part from this, students who undertake science stream in Form Four can opt to do Additional Mathematics. They also learn to understand and use measures of central tendency both for grouped and ungrouped data and solve related problems. One of the suggested activities at this level is that students collect real life situation data and investigate measures of central tendency. Student learn to determine the effect of mean, median, and mode for a set of data when each data is changed uniformly, extreme values exist and when certain data is added or removed. Students also learn to determine the most suitable measure of centre for a given data (Integrated Curriculum for Secondary Schools Curriculum Specifications Additional Mathematics Form Four, 2006).

Summary. Measures of central tendency hold a significant portion in the learning area of statistics under the Malaysian mathematics curriculum. Careful and detailed review of the curriculum indicates that the document embraces and supports the enhancement of statistical literacy through the emphasis given to the context of real-world by incorporating real life situations in the teaching and learning activities.

Although the curriculum supports the enhancement of statistical literacy but the question is whether our pre-service mathematics teachers have the subject matter knowledge of measures of central tendency involved in the enhancement of statistical literacy.

The Nature of Statistical Literacy

People's awareness of the application of statistical understandings in their daily lives have acknowledged and emphasized statistical literacy in statistics education. This was seen in the early 1990s when Katherine K. Wallman offered her perspective in the Presidential Address at the 1992 annual meeting of the American Statistical Association in Boston, Massachusetts.

According to Wallman (1993), "statistical literacy is the ability to understand and critically evaluate statistical results that permeate our daily lives-coupled with the ability to appreciate the contributions that statistical thinking can make in public and private, professional and personal decisions" (p. 1).

Wallman emphasized that citizens of quantitative world are users of statistics rather than creators and in order for them to properly function as citizens; they must possess the ability of statistical literacy for decision-making in their lives. Wallman also believed that the schools play a crucial role in producing graduates whom are knowledgeable statistics users.

In 1997, Watson presented a framework of statistical literacy that comprised of three tiers: basic understanding of probabilistic and statistical terminology; an understanding of statistical language and concepts when they are embedded in the context of wider social discussion; and a questioning attitude one can assume when applying concepts to contradict claims made without proper statistical foundation.

After almost a decade, Iddo Gal went further with his definition and proposed the two interrelated components that contribute to educated adults' statistical literacy:

> (a) people's ability to interpret and critically evaluate statistical information, data-related arguments, or stochastic phenomena, which they may encounter in diverse contexts, and when relevant.
> (b) their ability to discuss or communicate their reactions to such statistical information, such as their understanding of the meaning of the information, their opinions about the implications of this information, or their concerns regarding the acceptability of given conclusions. (Gal, 2002, p. 3)

These capabilities and behaviours according to Gal are supported by several interrelated knowledge bases and dispositions.

Gal believed that having the knowledge alone without the dispositional component is meaningless and would not lead to statistical literacy. He pointed out that when the optimal level of statistical literacy is reached, individuals will use the acquired knowledge bases and critical thinking skills and apply them on statistical information that they encounter in everyday life. A statistically literate adult should own a willingness to criticize statistical messages and this disposition emerges when the adult believes in his capabilities, has an interest and readiness to engage in statistical thinking.

Gal also mentioned that statistical literate action can be both overt and hidden because the action can involve internal mental processes, such as thinking about the meaning of something read, or raising in one's mind the critical questions related to it and reflecting about them. However, for any form of action to take place and be sustained, certain dispositions are needed and be activated (Gal, 2002, p. 18).

Thus, both Wallman and Gal have similar understanding of statistical literacy where statistical literacy is seen as a person's ability. However, Gal (2002) was more clear, precise and complete with his definition. He went further by elaborating on the types of statistical information, what to do with such information when encountered and finally proposed a model for the development of statistical literacy.

Recent studies however, have elaborated on the understanding of statistical literacy that included skills. One study that viewed statistical literacy as a set of skills was by Garfield, delMas and Chance (2003). They believed that statistical literacy is not only to comprehend, interpret, critically evaluate, and react to statistical messages but also is a set of skills used to organize, construct and display statistical information.

Similarly, Garfield and Ben-Zvi (2004) viewed statistical literacy as a set of skills that students should acquire from any statistics program. They mentioned that the skill to properly evaluate evidence (data) and claims based on data is an important skill that all students should learn as part of their school experiences.

On the other hand, statistical literacy is viewed as a mode of connectivity between knowledge and real life. Watson (2006) believed that statistical literacy is the "meeting point of the chance and data curriculum and the everyday world, where encounters involve unrehearsed contexts and spontaneous decision-making based on the ability to apply statistical tools, general contextual knowledge, and critical literacy skills" (p. 11).

Summary. Although, there were many interpretations given to statistical literacy but when statistical literacy is viewed as part of the description of people's ability of goal-oriented behavior in a specific domain, a broad cluster including factual knowledge, certain formal and informal skills, beliefs, habits of mind, or attitudes, as well as general awareness and a critical perspective are all involved (Gal, 2002).

However, in the present study, statistical literacy is viewed as the ability to utilise statistics in daily life situations. Specifically, the present study had defined statistical literacy as the ability to relate central tendency concepts to real life situations.

Ideas of Measures of Central Tendency Involved in the Enhancement of Statistical Literacy

Ideas of measures of central tendency taught in schools should not be confined to the routine computational ones but must be extended to the interpretation, critical evaluation of information and communication related to these measures. However, it is quite saddening to discover that students' experiences involving these measures are often reduced to routine computational ones (Shaughnessy, 2007). Therefore, the focus on the algorithm to produce a correct answer for every "average" problem needs to be replaced with a focus on context to make meaningful summary comments about data sets using suitable measures of centre (Watson, 2006, pp. 98 – 99).

Watson (2006) a prominent researcher in field of statistical literacy presented the ideas of measures of central tendency that should be included in the enhancement of statistical literacy; measures of central tendency in context, measures of central tendency and bias, measures of central tendency and problem solving, and measures of central tendency and inference.

Measures of central tendency with reference to context – The idea of average in statistical literacy can be associated with all the three measures of central tendency (Watson, 2006, p. 99). Assumption that the arithmetic mean is the average all the time is not correct. Being able to relate algorithm to concrete situations such as in the presentation of the average to two decimal places or the idea of mean which does not preserve numerical set (Batanero, Cobo, & Diaz, 2003; Watson, 2006) are considered important in relation to statistical literacy.

Apart from this, knowing typical behaviour when percentage is used should indicate that the average is referring to the mode. Reports based on human population should give rise to the awareness of the shape of the distribution and the presence of outliers. The averages in such reports might indicate that the median might have been used (Watson, 2006, p. 99).

Relating the average with the associated data set from which it is constructed also reflects statistical literacy (Mokros & Russell, 1995). Statistical literate readers of data when they see an average value; they think of the different distributions that these indicators could represent.

Making sensible decisions on the possible data set and how average can be obtained from the data set will provide some valuable insights on the representativeness of average and the appreciation of average in reducing the complexities of the data. Representativeness of average as a typical behaviour, "middle value", a balance point, a fair share or an algorithm (Mokros & Russell, 1995; Strauss & Bichler, 1988; Watson, 2006) are all considered important when statistical literacy is concerned.

Measures of central tendency in handling bias - Another important link to the ideas associated to measures of central tendency in relation to statistical literacy is the discussion on extreme values and which of the three measures of centre is least likely to be biased by these extreme values (Watson, 2006, p. 110). Understanding the distinctive features of the arithmetic mean and median can lead to an appreciation on why it may be important to use one or the other (Watson, 2006, p. 110).

Measures of central tendency in problem-solving - Deconstructing of mean can lead to a consideration of variation in the data set it represents. The representative nature of the mean can be considered in making sense of why data sets need to be combined to obtain a weighted mean. Although, algorithms are necessary but understanding the connections of the algorithms with the purposes of using summary statistics is equally important (Watson, 2006, p. 113). In order to solve weighted mean problems, the knowledge of applying the algorithm for the arithmetic mean in both the "forward" and "backward" direction is essential (Watson, 2006, p. 114).

Measures of central tendency in making inference - Measures of central tendency in reducing the data to single value(s), leads to inference stage of statistical investigations. One of the goals of statistical literacy is recognising measures of central tendency as significant tools for comparisons or judgements about data sets without being specifically told to use (Watson, 2006, pp. 118-119). In addition to this, average as "signal in noise" takes on the perspective that each observation is an estimate of the unknown but specific value (Konold & Pollatsek, 2004). The mean should be considered as a suitable measure and should be viewed as a meaningful real-world concept (Leavy & O'Loughlin, 2006, Pollatsek, Lima & Well, 1981) and not just a calculated measure. In any reporting of average, questions concerning range, outliers, spread, and shape should arise in the mind of any statistically literate individuals (Watson, 2006).

Summary. The above section presents the ideas of measures of central tendency involved in the enhancement of statistical literacy. Therefore, the present study had utilised these ideas in the development of the clinical interview tasks.

Learning Mathematics for Teaching: The Coding Rubric

The Learning Mathematics for Teaching (LMT) project was a project initiated due to the growing interest in the field to study the mathematical quality of teachers' practice. The project developed an instrument to quantify the quality of the mathematics in instruction that covered wide aspects of mathematical characteristics such as accurate use of mathematical language, the avoidance of mathematical errors, the provisions of mathematical explanations, the connection of classroom work to important mathematical ideas and so.

The argument behind the whole idea of this project was that constellation of lesson characteristics is not only a key but also the most important indication of a teacher's mathematical knowledge for teaching. Therefore, the instrument developed needed to be largely distinct from measure of teachers' pedagogical choices which have been the frequent focus of observational rubrics.

The project developed a measure that quantifies the quality of mathematics appearing in instruction regardless whatever teaching method. The development effort started with an attempt to validate pencil and paper measures of teachers' mathematical knowledge for teaching (MKT) (LMT, 2006). However, a serious need was seen on having a detailed rubric to map how mathematical knowledge for teaching might appear in practice, in which this became a measure development project by itself.

The project investigated on the mathematical knowledge needed for teaching along with how such knowledge developed as a result of teachers' experience and professional learning. As part of this work, the analysis was done not only based on whether teachers can solve the problems that they teach but also on how they work through the tasks which are unique to teaching. However, there were critiques on teachers' performance on pencil and paper assessments as to predict in class performance in which drew the attention of the project. Thus, they designed their validation work of their instrument.

The research project sought and identified the mathematical knowledge teachers drew upon and used during instruction. Furthermore, the project also focused on the development of methods for accurately reporting teachers' grasp of this knowledge. By having the above as the goals, the researchers worked on developing the coding rubric. A total of three phases were involved in the efforts to develop the codes: first was on mapping the terrain; second was on the use of the literature to support the existing codes that emerged as well suggested addition codes; and in third phase involved refine the codes and developed the glossary of standard procedures and definitions.

In the development of the video and actual coding process, the researchers 'blind' themselves to teachers' scores on the pencil and paper measure to avoid any influence in their coding scheme. In the mapping the terrain phase, the researchers had three goals for coding: 1) to track on the mathematical knowledge that appears in teaching, including agility or fluency in its use; 2) to watch for places where the teachers encounter mathematical difficulties and 3) to develop knowledge of the mathematical issues and problems that arise in teaching. The researchers wanted their codes to reflect both positive uses of mathematical knowledge as well as difficulties or mistakes.

Short segments of video tapes on how teachers' mathematics knowledge appeared in the lessons were watched and later discussions were carried on how the ideas might be translated into codes were done. After several meetings, the researchers developed four main categories of codes corresponding to the following questions:

1) What is the teacher's command of the mathematical terrain of this lesson?
2) How does the teacher know and use mathematical knowledge in dealing with students?
3) How does the teacher know and use mathematical knowledge in using the curriculum?
4) How does the teacher know and use mathematical knowledge for teaching equitably?

The above four categories remained constant throughout and became the foundation for the subsequent sections in which the researchers sought to identify finer grain elements in mathematical knowledge for teaching. The data from video recording suggested categories less often described in the existing literature. Each of the initial codes was originally measured using the five-point Likert Scale. The initial codes were revised as the researchers attempted to use them to code more video records of teaching.

The process of reviewing the videotapes revealed nuances and even new categories. As the researchers coded the data, they identified three elements of the mathematical knowledge in teaching; the presence of the mathematical description, explanation and justification. However, they found that defining the boundaries between these three elements of instruction proved difficult.

A panel of researchers from various background such as research mathematicians, former elementary teachers, mathematics educators and those with no formal mathematical or education background made the final coding that reflected the following distinction:

1) Mathematical description (of steps): Teacher's directing of mathematical descriptions (by self or co-produced with students) provides clear characterizations of the steps of a mathematical procedure or a process (e.g., a word problem). Does not necessarily address the meaning or reason for these steps. Code I for incomplete or unclear attempts.

2) Mathematical explanations – giving mathematical meaning to ideas or procedures: Teacher's directing of explanations (by self or co-produced with students) includes attention to the meaning of steps or ideas. Does not necessarily provide mathematical justification. Code I for incomplete and unclear attempts.

3) Mathematical justifications: Teacher's directing of explanations (by self or co-produced with students) include deductive reasoning about why a procedure works or why something is true or valid in general.

In the second phase, the researchers worked on checking the emergent coding scheme with the existing literature on mathematical knowledge for teaching. In this phase, they developed new codes as well as revised and refined existing codes.

At the development of a protocol for coding and record keeping phase, the researchers decided on the length of the lesson segments that they would code. They decided on a 5-minute segments. The analysis was done by aggregating codes to the lesson level and then to the teacher level but their ultimate score was on teacher level rather than lesson or segment level.

However, even after breaking the lessons into 5-minute segments, the researchers ran into difficulty designing a method for recording the observations. They initial used Likert scale coding but realized that Likert scale required too many subtle distinctions to reach reliability on the codes like "mathematical explanations" and "technical language". Thus, the researchers shifted the coding based on whether the actions indicated by the code were "present" (P) or "not present" (NP). However, they also realized that this also missed two important aspects which are whether elements that were present were also acceptable mathematically, or if the elements contained errors; and whether elements that were present should have occurred in order for instruction to reasonably proceed.

Therefore, to capture this difference, the researchers added an additional component of "appropriate" (A) and "inappropriate" (I) to the "present" (P) and "not present" (NP) coding options. This created four possible options for the codes in tables: Present and appropriate (P-A), present and inappropriate (P-I), not present and appropriate (NP-A), and not present and inappropriate (NP-I). Figure 2.6 shows a coding decision tree that represented the thought process.

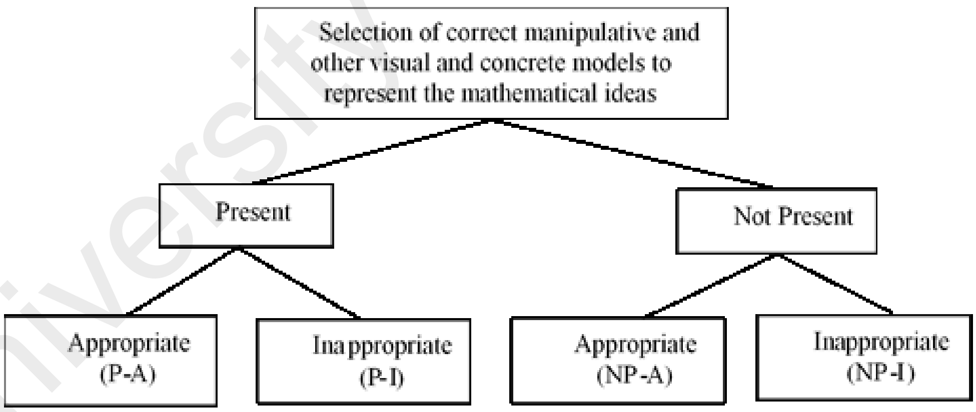


Figure 2.6. Coding decision tree

Summary. This section presents the overview of the coding thought process adapted in this study to answer the second research question related to the levels of subject matter knowledge of measures of central tendency of the pre-service mathematics teachers.

Chapter 3 Methodology

Introduction

This chapter discusses the methodology of the present study. The chapter consists of nine main sections: research design, selection of subjects, data collection, administration of the interview, instrumentation, content validity, pilot study, data analysis, and trustworthiness.

Research Design

The nature of the research problem and research questions of the present study, which sought to investigate and understand the pre-service mathematics teachers' subject matter knowledge of measures of central tendency in which little is known especially in the Malaysian context, influenced the researcher's decision that the present study is of a qualitative nature (Creswell, 2008).

The present study was guided by the following characteristics of a qualitative research; the goal of eliciting an understanding on the nature of the subject matter knowledge of measures of central tendency held by our Malaysian pre-service mathematics teachers with the researcher as the primary instrument in the data collection and analysis phases. The study also carried an inductive orientation towards the data analysis in order to obtain findings which are richly descriptive (Merriam, 1998)

Following these characteristics of a qualitative research, the present study was further aimed to gain an in-depth understanding on the nature of pre-service mathematics teachers' subject matter knowledge of measures of central tendency. The present study was interested in discovering the types of subject matter knowledge of measures of central tendency and the levels of subject matter knowledge of measures of central tendency held by the pre-service mathematics teachers.

Therefore, the case study design was found to be the most suitable design for the present study because the interest of the present study was on the process rather than outcomes, in the context rather than a specific variable, in discovery rather than confirmation. These insights are hoped to influence the practice and future research (Merriam, 1998, p. 19). In fact, several researchers have employed case study research design to study Malaysian students, pre-service teachers, and lecturers (Chew, 2007; Sharifah Norul Akmar, 1997; Wun, 2010).

The present study was also bounded by qualitative case studies' characteristics: (a) "particularistic – focus on a particular situation, event, program, or phenomenon; (b) descriptive – the end product is a rich, 'thick' description of the phenomenon under study; (c) heuristic – illuminate the reader's understanding of the phenomenon under study" (Merriam, 1998, pp. 29 – 30).

In terms of "particularistic", the present study focused on how the six pre-service mathematics teachers attempted the tasks related to the measures of central tendency. The present study examined pre-service mathematics teachers' instances when confronting the tasks but illuminated a general problem related to the subject matter knowledge of measures of central tendency.

In terms of "descriptive", the present study provided a rich, "thick" description of pre-service mathematics teachers' subject matter knowledge of measures of central tendency. The present study also illuminated the reader's understanding of Malaysian pre-service mathematics teachers' subject matter knowledge and levels of subject matter knowledge of measures of central tendency (refers to the characteristic on "heuristic").

Selection of Subjects

After deciding on the most appropriate research design, the next step is to select the unit of analysis or sample. Nonprobability sampling is the method of choice for most qualitative research because generalization is not a goal of qualitative research (Merriam, 1998, p. 61). This applied to the present study which is not aimed to generalize its results.

The purpose of the present study was to discover, understand and gain insight pre-service mathematics teachers' subject matter knowledge of measures of central tendency. The researcher selected a sample from which the most can be learned (Merriam, 1998, Patton, 1990). Therefore, in the present study, the researcher employed the non-probabilistic – purposeful sampling.

Merriam (1998) mentions that to begin purposeful sampling, the essential selection criteria in choosing the people or sites must be determined first. The criteria determined for the purposeful sampling must "reflect the purpose of the study and guide in the identification of information-rich cases" (Merriam, 1998, pp. 61-62).

In order to determine the number of subjects, Merriam (1998) mentioned that "The more cases included in a study, and the greater the variation across the cases, the more compelling an interpretation is likely to be" (p. 40). However, Patton (1990) recommends that specifying a minimum sample size "based on expected reasonable coverage of the phenomenon given the purpose of the study" (p. 186).

Thus, following the above guidelines mentioned by Merriam (1998) and Patton (1990), the following criteria were used to select the sample for the present study: (a) pre-service mathematics teachers enrolled in a 4-year Bachelor of Science with Education (B.Sc.Ed.) program in a public university in Peninsula Malaysia who majored or minored in mathematics, (b) they have successfully completed two years of the teacher education program and at the point of the data collection phase, they were either in their third or final year of program. The rationale behind these criteria is that by the end of the second year of the teacher education program, the subjects would have covered all the content related to statistics, (c) they agreed to serve as the participant for the present study, (d) they agreed to be audio recorded and video recorded during the clinical interview sessions, and (e) they agreed to complete the three clinical interview sessions. However, the subjects were informed that they have the freedom to withdraw from the study at any stage on their own will.

The selection of the subjects (sample) began when the researcher sought permission from the B.Sc.Ed. program course coordinator to conduct the study. The purpose and the nature of the study was explained to the course coordinator in order to gain permission and to collect background information data from the pre-service mathematics teachers.

The researcher collected the background information of the pre-service mathematics teachers using the Participant Background Information Form (see Appendix A). The Participant Background Information Form was used to collect pre-service mathematics teachers' demographic information, contact information, national examinations' mathematics results, information on their major and minor of the teacher education program, university content courses taken and the grades obtained for the respective courses, the cumulative grade point average (CGPA) secured, and teaching experiences (if any).

Before the researcher collected the background information of the pre-service mathematics teachers, the researcher explained the purpose and the nature of the study. The pre-service mathematics teachers were assured that the information that they provided in the Participant Background Information Form and the information gathered from the three clinical interview sessions will be kept confidential and will only be used for the purpose of the present study. They were assured that the present study would not affect in any way their performances in the teacher education program.

Therefore, based on all the above criteria, the present study collected data from six subjects (sample). The researcher selected three pre-service mathematics teachers who majored in mathematics and three pre-service mathematics teachers who minored in mathematics for the purpose of this study.

Based on the Participant Background Information Form, the researcher selected subjects that had a variety of major-minor combination, ethnicity, gender and academic achievements. This was believed to provide good variation across the cases so that the interpretation of the data would be more convincing (Merriam, 1998). The researcher ensured that the selection was done as random as possible. The selected pre-service mathematics teachers were also required to provide a written consent using Subject Informed Consent Form (see Appendix B).

Data Collection

Data for the present study was collected using the clinical interview technique in general. Clinical interview technique was found to be the most appropriate technique for the present study because it served the main purpose of obtaining a special kind of information by finding out what is "in and on someone else's mind" (Patton, 1990, p. 278).

According to Patton (1990):

> We interview people to find out from them those things we cannot directly observe. We cannot observe feelings, thoughts, and intentions. We cannot observe behaviors that took place at some previous point in time. We cannot observe how people have organized the world and the meanings they attach to what goes on in the world. We have to ask people questions about those things. The purpose of interviewing, then, is to allow us to enter into the other person's perspective. (p. 196)

Interview is also regarded as the best technique to use when conducting intensive case studies of a few selected subjects (Merriam, 1998, p. 72). Therefore, it served the purpose of the present study. In fact, numerous studies have used interviews as the main source of data collection (Ball, 1988; Fatimah, 1997; Nik Azis, 1987; Sharifah Norul Akmar, 1997; Wun, 2010).

However, in the present study the clinical interview was task based. Clinical task based interview gained its origin from clinical interview. This form of clinical interview has been used by qualitative researchers in mathematics education to gain knowledge about an individual or a group of students' existing or developing mathematical knowledge and behaviors (Maher & Sigley, 2014).

Thus, the present study used clinical task based interview to find out pre-service mathematics teachers' nature of the subject matter knowledge of measures of central tendency because how pre-service mathematics teachers have organized their subject matter knowledge of measures of central tendency and the meanings of the measures are things that cannot be directly observed. It can only be known through asking questions and entering into the subject's perspective by observing his or her actions while engaged in the mathematical task. During the clinical task based interview, the subject would not only interact with the researcher but also with task environment which is designed carefully for the purpose of the interview (Goldin, 2000).

In the clinical task based interview, the questioning and answering procedure was flexible, open, intensive, elicit, explore, and require the subjects to provide explanation about their actions to enable the researcher to gain understanding of their subject matter knowledge of measures of central tendency. The questioning and assessment during the course of the interview was used to elicit pre-service mathematics teachers' subject matter knowledge of measures of central tendency.

The communication was focused on a task that allowed the subject to display his or her behavior about the task that can be inferred. The researcher placed emphasis on the role of language and the importance of clarification of the meaning by asking questions and seeking explanation from the subjects about their actions (Hunting, 1997; Hunting & Doig, 1997).

The context of one to one interaction was used to make direct observation on the subject's behaviors when he or she attempted to answer or solve the tasks. The researcher treated each subject differently whereby the researcher thrived on the subject's individual variation. Based on the pre-service mathematics teachers' subject matter knowledge of measures of central tendency, the researcher later determined the pre-service mathematics teachers' levels (low, medium, high) of subject matter knowledge of measures of central tendency.

The materials that were collected for the analysis consisted of audio recordings, video recordings of clinical interviews, subject's notes in the task sheets, and researcher's notes during the interviews. The audio recordings and video recordings were verbatim transcribed into written form and the transcriptions which included both verbal and nonverbal interaction between the researcher and the subject served as the raw data for the present study.

Administration of the Interview

Each clinical interview session was conducted in the Discussion Room A in the main library at a public university in Peninsula Malaysia. The reason for choosing this venue was because all pre-service mathematics teachers who were interviewed were familiar with this room. The room was conducive which allowed the subjects to be more relaxed and comfortable with the surrounding environment.

The physical setting for each interview consisted of a table with two chairs, an audio recorder and a video recorder. Each interview was recorded using audio and video recorders. Blank papers, a pen, and a calculator were accessible to the pre-service mathematics teachers throughout the interviews.

Each of the six subjects was individually interviewed for a duration between 30 to 40 minutes on three separate occasions in a period of twenty-four weeks. Each session involved four to seven tasks (see Appendixes D-F).

The total time for each interview varied from subject to subject depending on the amount of time taken to complete the tasks. Each interview session started with the researcher posing a task to the subject. The subject responded to the task. The researcher probed the subject based on his or her response to the task by either presenting the task in a different way or by posing a new task. This was necessary to gain valuable insights from the subjects. Due to this, each interview was different from one another because each subject is unique.

Instrumentation

A total of fifteen tasks covering four constructs namely; measures of central tendency with reference to context, measures of central tendency in handling bias, measures of central tendency in problem solving, and measures of central tendency in making inference were planned and used in the present study (see Appendices D-F). These tasks were based on the theoretical perspective and the conceptual framework of the study.

The following table illustrates the description of the tasks used in the present study.

Table 3.1

Description of the tasks

Task	Description	Component
Task 1 (Mean)	Task 1 was adopted from Watson and Moritz (2000). The general term average can mask any one of the three measures of central tendency. In this task, the average is the mean is indicated through its numerical form in relation to the associated data set (Watson, 2006). The task explores what sense the subjects can make from a decimal average and if they are able to relate to the mean. The task explores if the subjects can relate algorithm to concrete situation and the use of algorithm can produce a value that does not preserve the numerical set of the associated data set. This is considered important in relation to statistical literacy because this provides the knowledge that an average is the mean if it is presented in a form that does not preserve the numerical set of its associated data set.	Measures of central tendency with reference to context
Task 2 (Mode)	Task 2 was developed by the researcher. Task 2 explores the subjects' knowledge of the mode. Mode can represent typical behaviour in which this can be used as a quick method of reporting central tendency and is described as the only acceptable measure of central tendency for categorical variable. In relation to statistical literacy, the reporting of an average can also be based on this characteristic of the mode.	Measures of central tendency with reference to context
Task 3 (Mode)	Task 3 was adapted from Konold and Garfield (1992). The task explores the subjects' knowledge of the mode that can represent typical behaviour in which this can be used as a quick method of reporting central tendency and is described as the only acceptable measure of central tendency for categorical variable. The modification made in the original task was that the data in the original task was presented as numerical but for the purpose of this study, the data was arranged in tabular form. This was done to create a situation for the identification of mode when the data is presented in tabular form and to give variety in the data over the devised tasks. The question in the original task was also restructured because the word "typical" was not common to the subjects which was observed during the pilot interview.	Measures of central tendency with reference to context
Task 4 (Median)	Task 4 was adapted from Watson (2006). The task explores the subjects' interpretation of the term average given in a context that had actually masked the use of median. In this task, the subjects' familiarity of extreme values and the basis used to choose the suitable measure of central tendency to summarise the associated data set was explored. The modification made to the original task was that the situation was changed to the Malaysian context. The researcher also discovered from the pilot interview that the subjects failed to interpret the given average as the median. Thus, the researcher modified the task by adding information on an outlier group in the task situation. With such information, the researcher wanted to explore if the subjects are aware that in the presence of outliers, the median is the best measure to represent the data and that the average was based on the median.	Measures of central tendency with reference to context
Task 5 (Median)	Task 5 was adopted from Batanero, Merino and Diaz (2003). Task 5 is a straightforward task that explores the idea of median and the biasing effect of an outlier on the median. In	Measures of central tendency

	the task, the subjects' justification on the most suitable measure between the median and the mean was explored.	with reference to context
Task 6 and Task 7	Task 6 and Task 7 were adapted from Leavy and O'Loughlin (2006). The original tasks were taken from Mokros and Russell (1995). Both the tasks were changed to the Malaysian context where by the average price of 27p was changed to RM 27. These tasks provide further insight on the subjects' understandings of the relationship between the average and the data set from which the average was constructed. Thus, further provided valuable insights into their understanding of representativeness of the average whether it is a typical value, a middle value, a balance point, a fair share or an algorithm which contribute towards their understanding of average in context. According to Mokros and Russell (1995): Construction problems are very revealing of students' understanding of the relationship between data and average. This kind of task also explores what statistically literate readers of data do when they see a mean or a median; if they think of different distributions that could be represented by this indicator". (p. 23)	Measures of central tendency with reference to context
Task 8 and Task 9	Task 8 was adapted from Konold and Pollatsek (2004). The original task was from Konold and Garfield (1992). Whereas, Task 9 was developed by the researcher. The modification done to the original Task 8 was to the questions that followed the task. The researcher devised her own set of questions based on the context that needed to be investigated. Task 8 and Task 9 are used to explore how the subjects use measures of central tendency to summarise data sets. These tasks explore the basis used to choose the suitable measure of central tendency by looking at the advantages and disadvantages of the three measures of central tendency in making a conclusion or an inference on the most suitable measure of central tendency to represent the data set. The tasks were also used to explore on whether the subjects are able to deal with the issue of bias in the presence of an outlier measurement in a given data set.	Measure of central tendency in handling bias
Task 10	Task 10 was adopted from Pollatsek, Lima and Well (1981). Task 10 is a weighted mean problem which is used to explore the subjects' ability to combine the algorithmic numerical skills with the appreciation of the context within which they are used. However, Task 10 does not explicitly state to solve for the weighted mean. Mean has a real-world referent that constraints the choice of scores that can be entered in a computational formula (Pollatsek, Lima, and Well, 1981). Solving a weighted mean problem involves in making sense of why one would want to combine data sets to obtain a weighted mean. Although it is necessary to know the algorithm but it is also important to know their connections with the purposes of using summary statistics. These are considered as the more technical and sophisticated goals of statistical literacy.	Measures of central tendency in problem-solving
Task 11	Task 11 was also adopted from Pollatsek, Lima and Well (1981). Task 11 is also a weighted mean problem similar to Task 10 but Task 11 explicitly states to solve for the weighted average.	Measures of central tendency in problem-solving

Task 12	Task 12 was adapted from McGatha, Cobb, and McClain (2002). The researcher modified the original task to Malaysian context and changed some of the scores to create a situation for the existence of all three measures of central tendency. The researcher also created a situation where by both players in the task have equal scores in order to create two equal sized data sets. Therefore, Task 12 is used to explore how the subjects compare two equal sized data sets presented in numerical form because one of the goals of statistical literacy is to have a range of statistical tools in which measures of central tendency are one of the significant ones that can be used in the comparison and judgement about data sets without specifically being told.	Measures of central tendency in making inference
Task 13	Task 13 was also adapted from McGatha, Cobb, and McClain (2002). The researcher modified the original task to Malaysian context and changed some of the scores to create a situation for the existence of all three measures of central tendency. This task is used to explore how the subjects compare two unequal sized data sets. The mean which carries a function for its use that takes into account every score has an advantage in terms of the measure of central tendency in representing all of the data. Therefore, the mean is a meaningful real-world concept that can be used to compare two unequal sized data sets. Nevertheless, the task itself does give allowance for the subjects to use other measures of central tendency apart from the mean in which the task can be further used to explore on the subjects' explanation or justification for using those measures.	Measures of central tendency in making inference
Task 14 and Task 15	Task 14 and Task 15 was adapted from Watson and Moritz (1999). Both the tasks were modified to the Malaysian context and the researcher devised the task based on the original graphs in the original tasks. "Initial experiences with inference can be linked with predictions from data sets presented in graphical form…" (Watson, 2006, p. 190). Task 14 and Task 15 were used to explore how subjects compare two equal and unequal data sets presented in graphical form without the task explicitly stating to use the measures of central tendency.	Measures of central tendency in making inference

The tasks and the standard probes for the clinical interviews are given in Appendices D-F. Supplementary probes were generated in the course of the clinical interview and whenever necessary in order to obtain a clearer picture about what the subject was thinking. Investigation into pre-service mathematics teachers' subject matter knowledge of measures of central tendency required efforts at probing as deeply as possible by asking them to justify their responses. This gave the researcher, an in-depth view of the nature of the subject matter knowledge of measures of central tendency in the respective tasks of the clinical interviews.

Table 3.2

Distribution of the subject matter knowledge of measures of central tendency clinical interview tasks based on the four components of investigation

Construct	Interview 1							Interview 2				Interview 3			
	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Task 7	Task 8	Task 9	Task 10	Task 11	Task 12	Task 13	Task 14	Task 15
Measures of central tendency with reference to context	X	X	X	X	X	X	X								
Measures of central tendency in handling bias								X	X						
Measures of central tendency in problem solving										X	X				
Measures of central tendency in making inference												X	X	X	X

Pre-pilot Interview

The researcher conducted a pre-pilot interview in June 2014 to validate the written instruments in terms of task relevance, content coverage, and language appropriateness before the proposal of the present study was vetted. The pre-pilot interview helped the researcher to determine any confusing questions, questions that needed rewording, questions that would not yield data, and questions that needed to be included in the present study (Merriam, 2009).

From the pre-pilot interview, the researcher discovered that some of the tasks needed to be reworded because the subject had problem understanding the questions. The researcher realized that she had to preset some probing questions to yield the necessary data.

The pre-pilot interview also helped the researcher to sequence the tasks according to the interviews so that the subjects would not get any hint from the earlier interviews and tasks on the topic of investigation. For instance, Task 12 and Task 13 that involved the investigation of the subject's knowledge of the mean as a functional measure without the tasks explicitly stating it, therefore sequencing these tasks separately in a different interview session would yield better data on the subject's knowledge of the mean as a functional measure in real world situation.

The verbatim transcript of the pre-pilot interview gave the researcher a strategy to look for places where probes should have followed up so that to maximize an opportunity to gain more information.

Content Validity

Content validity is done to show that the questions or tasks on a test or instrument is representative of some defined universe or domain of content (Hogan, 2007; McMillan, 2001; Popham, 2000).

The content validity of the present study was achieved using experts' reviews on the tasks. The panel of experts consisted of two university lectures. The university lectures specialized in mathematics and statistics education. The panel of experts were asked to judge the relevance and content coverage of the tasks that represented measures of central tendency in the four constructs of investigation namely; with reference to context, in handling bias, in problem solving, and in making inference.

The clinical interview tasks, the description of the tasks, Task Relevance Judgement Form (refer to Appendix H) and Content Coverage Judgement Form (refer to Appendix G) were given to each of the panel of experts. The tasks were examined in terms of the tasks' relevance to the topic of measures of central tendency using the Task Relevance Judgement Form and the content coverage of the tasks to the topic of measures of central tendency using the Content Coverage Form.

The content validation of the tasks in the present study was done twice. The first round was done before the proposal of the present study was vetted in January 2015 and the second round of validation was done after the suggestions on the improvement of the tasks was forwarded by the proposal defense panel. The proposal defense panel suggested to include tasks of different form of representations like tabular and graphical, to cover more thoroughly on all three measures of central tendency and to incorporate tasks which are appropriate for clinical interviews. The researcher took note of these suggestions and improved on the tasks before validating the tasks with the experts in the second round of validation.

Table 3.3 shows the summary of the results on the task relevance judgement by each of the panel of experts and the mean value of the judgement on each task. Table 3.4 shows the summary of the results on content coverage judgement by each of the panel of experts. It was apparent that all the values of judgement assigned by the experts were at least 4 and the mean value of the judgement on each task were at least 4. Thus, it can be concluded that all the tasks were relevant and cover the topic of measures of central tendency.

The experts also gave some valuable suggestions which were taken note such as to replicate the tasks so that the pattern of responses from the subjects on each type of tasks can be confirmed which would add to the credibility of the findings. Suggestion was also given to check on the relevance of certain tasks in yielding the relevant data.

Table 3.3

Judgement of Task Relevance

Task	Judge 1	Judge 2	Mean
1	4	5	4.5
2	4	5	4.5
3	4	5	4.5
4	4	5	4.5
5	4	5	4.5
6	4	5	4.5
7	4	5	4.5
8	4	5	4.5
9	4	5	4.5
10	4	5	4.5
11	4	5	4.5
12	4	5	4.5
13	4	5	4.5
14	4	5	4.5
15	4	5	4.5

Judgement scale

5 – most relevant 4 – quite relevant 3 – relevant 2 – less relevant 1 – not relevant

Table 3.4

Judgement of Content Coverage

Judge	Mean
Judge 1	4
Judge 2	5

Judgement scale
5 – most comprehensive 4 – quite comprehensive 3 – comprehensive
2 – less comprehensive 1 – not comprehensive

Pilot Study

The first pilot study was conducted in November 2014 prior to the actual field of data collection. A total of 11 tasks were devised for the pilot study. The tasks were validated by three experts and a pre-pilot interview. The experts suggested to include tasks of similar types to see any variations in the subjects' responses and also to re-sequence some of the tasks so that the subjects would not get any hint on the focus of the investigation.

Some of the experts' suggestions and outcomes from the pre-pilot interview were taken note before the tasks were used for the pilot study. The pilot study involved one pre-service mathematics teacher from a public university in Peninsula Malaysia. The subject majored in mathematics.

"The key to getting good data from interviewing is to ask good questions; asking good questions takes practice. Pilot interviews are crucial for trying out your questions" (Merriam, 2009, p. 95). This was found relevant for the researcher. The pilot study allowed the researcher to familiarize with the procedures and techniques of conducting clinical interviews.

The pilot interview also helped the researcher gain a good practice on using verbal probing. The researcher gained practice in administering series of probe questions designed to elicit specific information from the subjects beyond what the subjects have provided.

Besides this, the verbatim transcript of the pre-pilot interview gave the researcher a strategy to look for places where probes should have followed. These probes were further practiced in the pilot interview. This was necessary to check if these probes could maximize an opportunity to gain more information and elicit the necessary data. The pilot interview allowed the researcher to estimate the needed time for each interview. The researcher also confirmed from the pilot study that the clinical interview tasks were ready to be used in the actual field of data collection.

However, a second pilot interview was carried out in March 2015 after the instrument and the tasks were improved based on the suggestions given by the proposal defense panel. The proposal defense panel suggested to include more tasks that included the concept of mean, mode and median. Suggestion was also given to include tasks of various representations such as numerical, graphical and tabular.

Thus, new tasks were added to the interview protocol. These tasks were validated by two experts and a second pilot interview was carried out to confirm if the tasks were ready for the actual data collection. The validation judgements related to the overall task relevance and content coverage for all the 15 tasks used in the actual data collection were reported in the Content Validity section.

The second pilot interview involved one pre-service mathematics teacher from a public university in Peninsula Malaysia. The subject minored in mathematics. From the second pilot interview, the researcher found that some of the new tasks needed rewording such as Task 3 where by the original task had the word “typical”. However, this word was uncommon to the subjects in the Malaysian context. Therefore, the researcher modified and restructured the questions that followed this task.

The researcher also discovered from the pilot interview that the subject failed to interpret the given average for a population which was based on the median in Task

4. Thus, the researcher modified the task by adding information on an outlier group in the task description. With such addition, the researcher further hoped to explore if the subject is aware that in the presence of an outlier group, the median is the best measure to represent the data related to a population and that the average has to be based on the median.

From the pilot interviews, the researcher confirmed that the tasks were ready to be used in the actual data collection because the tasks could yield the necessary data. The following excerpts describe part of the data from the second pilot interview related to Task 3 which explored the subject's idea of the mode. The excerpts also describe how the researcher carried out her interviewing skills.

The subject explained on how she obtained the mode; *"…we can see 22 as the mode, the highest number in our sample. So, 22 is the highest and it is the mode".*

The researcher probed further and asked on how the subject obtained the mode. The subject further clarified that the mode is the highest number *"It is clear that the highest number here (referring to Table D-1) is 22. So that is the mode"*

The researcher modified the task by substituting 22 to 6 in the table and asked the subject to identify the mode for this situation. The subject now mentioned; *"It is 6"* after probed why she mentioned *"because the highest number now is 6 so the mode is the highest so it is 6 now"*.

The researcher questioned again if there exist any other mode for the given task. The subject replied; *"no... no..., only one mode"*. The researcher noticed that the subject was very confident when she answered and had least doubt over her response.

The above excerpts showed that the subject's knowledge of the mode as the highest value or the biggest value. Based on all the pilot interviews and validation judgements from the experts, the tasks were ready for the actual data collection phase which was carried out in April 2015 till August 2015.

Data Analysis

The data analysis of the present study was conducted simultaneously with the data collection which is the basic principle of a qualitative research (Coffey & Atkinson, 1996). The analysis and interpretations of pre-service mathematics teachers' subject matter knowledge of measures of central tendency were based on the data collected using audio recordings, video recordings, the researcher's notes, and the subjects' work on the three clinical interviews conducted on each of the six selected subjects. The interview data were verbatim transcribed, coded or categorized, and analysed.

The data analysis process was categorized into four different stages. At stage one, the raw data from the audio and video recordings of the clinical interviews were verbatim transcribed into written form. The transcription also included the subjects' nonverbal behaviours which were recorded through a video and the analytic memos on the interaction between the subjects and the researcher.

At this stage, the researcher developed many ideas or concepts as possible without having any concern over how these ideas or concepts would be used later. The researcher also went through the subjects' work in the interview task sheets and developed as many ideas or concepts too. Apart from this, the interview data was transcribed at the end of each interview sessions and before the subsequent interview sessions took place, the researcher went through the transcriptions with the subjects. This was done to rule out the possibility of any misinterpretations, to further understand the data, and finally to ensure the validation of the transcriptions.

From stage one, the coding process preceded to stage two where the refinement and development of categories which stood out were done. At this stage, the identified categories were further refined which lead to the identification of specific categories in which these specific categories provided a protocol related to the description of the subjects' subject matter knowledge of measures of central tendency that answered research question one, the types of subject matter knowledge of measures of central tendency.

At the coding stage, the researcher developed a set of codes or categories using the ones that are predefined and the ones that emerged from the data. Coding rubrics for determining the subjects' subject matter knowledge of measures of central tendency were devised to facilitate the coding, categorization, and analysis processes of each subject's subject matter knowledge of measures of central tendency (refer to Appendix J). The researcher continued to write analytic memos and hold analytic meetings with her supervisors to reflect on the researcher's initial ideas and examine them in relation to new information and further pose ideas about developing relationships and patterns.

In order to answer research question two, each subjects' levels (low, medium, high) of subject matter knowledge of measures of central tendency was analysed using coding rubrics and the coding thought process which were adapted from the Learning Mathematics for Teaching (LMT) (2006) project (see Appendix I).

The following elaborates on the coding process. The researcher determined whether the statistical element was present (P) or not present (NP). If the statistical element was present (P), then she marked: (a) appropriate (A) if the pre-service mathematics teacher's use of the statistical element was appropriate, accurate, or correct; or marked (b) inappropriate (I) if the pre-service mathematics teacher's use of the statistical element was inappropriate, inaccurate, or incorrect. If the statistical element was not present (NP), then she marked (a) appropriate (A) if the absence of the statistical element seemed appropriate or not problematic; or marked (b) inappropriate (I) if the absence of the statistical element seemed inappropriate or problematic (i.e., the statistical element should have been present) (adapted from LMT, 2006).

A preservice mathematics teacher's level (low, medium, high) of subject matter knowledge of measures of central tendency for each of the construct was determined based on the percentage of appropriate statistical elements of the subject matter knowledge of measures of central tendency. The following illustrates the computation of the percentage of appropriate statistical elements of subject matter knowledge of measures of central tendency:

Percentage of appropriate statistical elements of subject matter knowledge of measures of central tendency was calculated based on the following formula = $\frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)}$ x 100%, where nPA, nPI, nNPA, and nNPI represented the number of codes that were coded as "present and appropriate (PA)", "present and inappropriate (PI)", "not present and appropriate (NPA)", and "not present and inappropriate (NPI)", respectively.

A preservice mathematics teacher's level (low, medium, high) of the overall subject matter knowledge of measures of central tendency was determined using the grading system used in the university where the data of the study was collected. In order to obtain a preservice mathematics teacher's level of subject matter knowledge of measures of central tendency, first the mean percentage of the subject matter knowledge for all four constructs namely; with reference to the context, in handling bias, in problem solving, and in making inference was calculated.

The following describes on how the mean percentage was calculated for all four constructs. Mean percentage of the subject matter knowledge of measures of central = $\frac{A+B+C+D}{4}$ % where A represented the percentage of subject matter knowledge with reference to context, B represented the percentage of subject matter knowledge in handling bias, C represented the percentage of subject matter knowledge in problem solving and D represented the percentage of subject matter knowledge in making inference.

If a preservice mathematics teacher achieved 80% and above as the mean percentage of the overall subject matter knowledge he or she was assigned a high level of subject matter knowledge of measures of central tendency. If a preservice mathematics teacher's achievement was 40% to less than 80%, he or she was assigned a medium level of subject matter knowledge of measures of central tendency. Finally, if a preservice mathematics teacher's achievement was less than 40%, he or she was assigned a low level of subject matter knowledge of measures of central tendency.

At stage three, case study for each subject was constructed based on the information from the written protocol. At this stage, analysis was carried out to describe each subject's behaviours in solving every tasks or problems. At stage four, cross analysis was conducted that was aimed to identify pattern of responses of subject matter knowledge of measures of central tendency held by the subjects. Based on these pattern responses, pre-service mathematics teachers' subject matter knowledge of measures of central tendency and level of the subject matter knowledge of measures of central tendency were summarised according to the emerged themes.

Trustworthiness

The trustworthiness of the present study considered the following four criteria: credibility, dependability, transferability and confirmability (Cuba, 1981). To ensure the credibility or internal validity of the findings, although the present study employed purposive sampling but this was done in a random manner so that to negate charges of researcher bias in the selection of subjects. This kind of sampling ensured that any "unknown influences" would be distributed evenly within the sample (Preece, 1994).

Apart from this, triangulation was also carried out using interview and follow-up interview data collected from tasks of similar types. The researcher also triangulated the data with other documents relevant to the phenomenon of the study such as the subject's notes in the interview task sheets (see Appendices D-F) and the researcher's notes.

Member checking or respondent validation was another strategy to ensure the credibility of the findings. This strategy is considered the most important strategy to rule out the possibility of misinterpreting the meaning of what the subjects say and do and the perspective they have on what is going on. This strategy is also considered an important way of identifying the researcher's own biases (Maxwell, 2005).

In the present study, the process involved checking with the subjects on what is "said" match the "intended" at the end of each data collection session and also taking the preliminary analysis back to the subjects and ask if the researcher's interpretation "ring true". The researcher's interpretations of the data were shared with the subjects so that the subjects were given the opportunity to discuss and clarify the interpretations. They were also given the opportunity to contribute new or additional perspective under the issue of study. The researcher utilised the Interview Transcript Review form (see Appendix C) to ensure this part was carried out in this study.

Peer examination or peer review was another strategy used to ensure the credibility of the findings of the present study. This involved asking knowledgeable peers to scan through the raw data and assess whether the findings are plausible based on the data.

To ensure the dependability or reliability of the findings of the present study, the processes of the study was reported in detail thereby enabling any future researcher to repeat the work if not necessarily to obtain the same results. Such coverage also highlighted the extent to which proper research practices have been followed. Apart from this, the researcher implemented a process of double coding where a set of data was coded, and then after a period of time the researcher returned to the same set of data and coded again. The results of the codes were compared (Krefting, 1991). This was done to ensure the consistency of the findings or "dependability" of the data.

Interviews and follow-up interviews that contained tasks of similar types was another way used to obtain consistent and dependable data. Moreover, the researcher engaged in frequent debriefing sessions with her supervisors to discuss on the approaches, test the researcher's interpretations and recognize her own biases to ensure the reliability of the findings.

Transferability or external validity was achieved by providing sufficient thick description to contextualize the study so that the readers were able to determine the extent to which their situations match the present research context, and, hence whether the findings of the present study can be transferred (Merriam, 2009). In the present study, the researcher conveyed to the reader the boundaries of the study such as the number of subjects involved, the data collection methods employed, the number and length of the data collection sessions and so which are important before any attempts at transference can be made.

In terms of confirmability, triangulation played an important role in the context to reduce the researcher bias in the present study. Triangulation was carried out using interview and follow-up interview data from tasks of similar types. Data obtained from other documents relevant to the phenomenon of the study such as the subject's notes in the interview task sheets or the researcher's notes was also used in the triangulation.

The trustworthiness of the present study also depended on the integrity of the researcher in carrying out the study in an ethical manner. The researcher got a written consent from each of the subjects and explained to them on the ethical principles such as their anonymity and data confidentiality using the Subject Informed Consent Form. The subjects were also explained that they have the right to withdraw from the study at any point. This was done to ensure that the subjects had some idea of what to expect from the interviews and increased the likelihood of honesty in the subjects.

Apart from this, the researcher also forwarded the proposal of the study to University Malaya Research Ethics Committee (UMREC) for the Research Ethics Clearance approval. The committee approved that the present study followed all the required University of Malaya Research Ethics Guidelines (see Appendix K).

Chapter 4 Findings of the Study

Introduction

This chapter provides the findings from the subjects' analysis and descriptions of preservice mathematics teachers' behaviours as they attempted to solve each of the fifteen tasks during the clinical interview sessions. The findings from the analysis and description of the behaviours of each subject, namely Lina, Bella, Harry, Amy, Tina, and Joan were presented in case studies in this chapter.

To answer research question one, findings of preservice mathematics teachers' subject matter knowledge of measures of central tendency were presented in four different constructs namely; with reference to the context, in handling bias, in problem solving, and in making inference that were emerged from the clinical interview. To answer research question two, findings of preservice mathematics teachers' level of subject matter knowledge in each of the four-different construct namely; with reference to the context, in handling bias, in problem solving, and in making inference were calculated as well as the overall levels of subject matter knowledge for each of the subjects that were obtained from the clinical interview.

Case Study One: Lina

Lina was 24 years 1-month old when she was interviewed. At the point of data collection, she was pursuing a 4-year Bachelor of Science with Education (B.Sc. Ed.) program at a public university. She majored and minored in mathematics and chemistry respectively. Lina obtained A1 in Mathematics and Additional Mathematics in her SPM examination (equivalent to O level examination). She also scored an A in Mathematics in the Matriculation examination.

At the time of data collection, Lina was in the final semester of fourth year studies. She attained 3.62 in the Cumulative Grade Point Average (CGPA) for her first three years of studies in the public university. She does not have any teaching experience prior to this interview. The following sections discuss the findings of Lina's subject matter knowledge of measures of central tendency in each of the four different constructs, namely with reference to context, in handling bias, in problem solving, and in making inference that were emerged from the clinical interview.

Types of subject matter knowledge.

Subject matter knowledge of measures of central tendency with reference to context. The subject matter knowledge of measures of central tendency with reference to the context deals with the aspects of measures of central tendency presented in the context of daily life situation. This construct is used to explore the ideas of measures of central tendency namely; terms, procedures, algorithm and justification related to the measures of central tendency in the context of daily life situations. A total of 7 tasks were constructed to explore this construct.

Task 1 (Mean). In Task 1, the given average is the mean which is indicated through its numerical form in relation to its associated data set. This task is used to explore if the subject could relate the idea of mean in real life situation and if the subject is aware that the use of the mean algorithm can produce an average value that does not preserve the numerical set of its associated data set.

Table 4.1

Task 1

Let's say you are watching TV, and you hear: "On average, Malaysian families have 2.3 children." How do you think the average was obtained? Can you show and explain using an example? How can the average be 2.3, and not a counting number like 1, 2, 3, or 4?

Lina's knowledge of the mean is average. This was clearly seen in Lina's response when she recalled the mean algorithm and utilised it to find the average. Furthermore, Lina was aware that a decimal average can be illogical given to the context of the data in which she knew is produced by the use of the mean algorithm.

Table 4.2

Excerpt L1

R: How do you think the average was obtained here? S: If according to the formula, then it is going to be the sum of children divided by the number of families that taken account. R: Do you think that there are any other ways that would have been used to find the average besides the one that you have mentioned? S: Not sure. I don't know. R: Can you show and explain using an example on how the average would have been obtained in this task? S: Something like that [writes the example in the space provided in the task sheet] R: Here the average is 2.3 and not a counting number like 1, 2, 3. S: To me it is not logic because we are talking about children which has to be a whole number like 1, 2, and so on. It is not supposed to be a decimal. R: Then how can there be a decimal value given in this task? S: To me, maybe they are not taking into account what they are considering. They just like, ahhh the average is sum of divide the number and that's it. They do not take into account that the counting {I think what the subject meant is the calculation} is for the number of children or the marks. For them whatever is the answer that's it. R: So, what are you trying to say here regarding the value 2.3? S: The value 2.3 is obtained straight away from the formula.

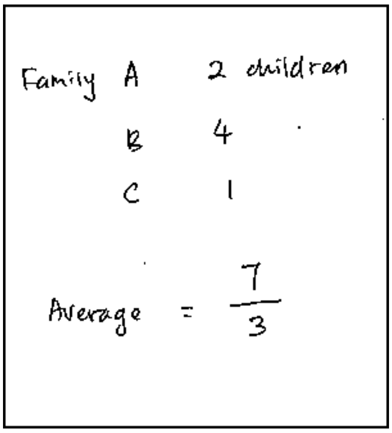


Figure 4.1. Entry L1

In Task 1, when Lina was asked on how the average was obtained, she mentioned the formula and elaborated as the sum of children divided by the number of families. Lina was not sure if there were other ways of finding the average besides the one that she mentioned. When Lina was asked to show, and explain on how the average would have been obtained in this task, she actually performed the arithmetic mean calculation which she wrote clearly in the task sheet as shown in Entry L1.

Lina expressed that the decimal value given to the average is not logic. She said that the ideal average should be a whole number. Nevertheless, when Lina was probed further on how the decimal value was obtained, she elaborated that the average was obtained using the mean calculation in which she expressed as "sum" and "divide". Lina claimed that in this situation the context of the data which was referring to the number of children was not considered. Instead the average was obtained straight from the use of the formula. The above Excerpt L1 illustrates Lina's points related to this.

Task 2 (Mode). In Task 2, the knowledge of the mode is explored. The mode can represent a typical behaviour in which this characteristic can be used as a quick method of reporting central tendency. In relation to statistical literacy, this characteristic of the mode can be used to report an average.

Table 4.3

Task 2

Information below shows female shoe sizes purchased in your shop.

4	5	5	5	6	7	6	6	7	4	4
5	5	5	7	7	7	8	9	5	5	4
6	7	8	9	5	5	6	7	5	4	5

How would you place the order for the shoe sizes in your shop?
If you want to choose one female shoe size, which shoe size would you choose? Why?
Can the chosen shoe size represent the female shoe size in your shop? Why?
What would you use as the average shoe size?
Why?

Lina utilised the idea of mode in daily life situation without the task explicitly stating so. She utilised the mode as a form of data representation but did not utilise the mode as a quick method of reporting an average.

Table 4.4

Excerpt L2

R: How would you place the order for the shoe sizes in your shop?
S: Depending on the stock and the frequent size that the customers wore.
R: How do you do that?
S: I will check the stock and then order based on the shoe size that is purchased the most and the rest follows.
R: If you want to choose one female shoe size, which shoe size would you choose?
S: five
R: Why?
S: Because… according this data, there are customers favour size 5 compared to the rest.
R: Can the chosen shoe size represent the female shoe size in the shop?
S: Can because the majority bought that.
R: What would you use as the average shoe size?
S: 5
R: Why 5?
S: There are two answers because of the majority and because of the formula we get 5 point something. So, it is 5.
R: You mentioned the formula, five point something. Can you explain further?
S: Ok, were to explain this. This is a group data because there are a quite a number of customers buying….. {} [sounded uncertain about the answer]
R: Just now you answered that because of majority and then you mentioned the formula. Can you explain further?
S: Umm… When we use the formula sum of fx over sum of x, so we group the data according to the size. We group it one by one, size 4, size 5 and get the frequency for each size so then total it up and divide.
R: What is the average here?
S: 5 point. I did not calculate. 5 point something. Should I calculate?
R: Maybe, if you can explain better if you have calculated the value.
S: 5.4
R: What would you use as the average shoe size?
S: Size 5
R: Why?
S: Because majority bought that size and is because of the average 5.4

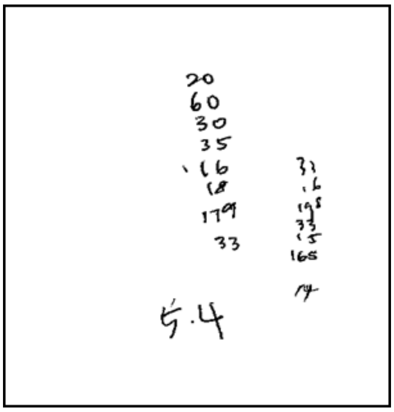


Figure 4.2. Entry L2

When Lina was asked on how she would place the order for the shoe sizes in her shop, she mentioned that she would place the order by looking at the stock and the frequent sizes that the customers wore. She explained that she will check the stock and then order based on the shoe size that is purchased the most and the rest of the orders will follow accordingly.

When Lina was asked to choose one female shoe size, she chose size five. She explained that this is due to size five was favoured by her customers. Furthermore, Lina mentioned that the chosen shoe size can represent the shoe size in her shop because “majority” bought that particular shoe size.

In addition, Lina mentioned that she would use five as the average shoe size. She based her decision on two different answers which are the “majority” bought that particular shoe size and the formula. Nevertheless, when Lina was probed to explain further, she looked rather uncertain on how to explain her answers earlier. After when, Lina explained in a rather unsure tone, first she recalled the formula to obtain the mean for grouped data. She mentioned as the sum of fx over sum of x. She elaborated that the data has to be grouped according to the size and the frequency for each group. Finally, she mentioned to total all up and divide.

However, her explanation was not very convincing which made the researcher to probe her with another question: what is the average? Now, Lina mentioned five point something but uttered that she did not calculate the value. In fact, she clarified if she should calculate. The researcher advised her if she felt that she could explain better by having the value calculated then maybe she should. After when, Lina performed the calculation and obtained the value as five point four. It was clear that “the formula” that Lina was referring to earlier was actually the mean algorithm because she had utilised the mean calculation to obtain the average as five point four which is shown in Entry L2.

Looking at Lina’s response, the researcher probed her again by repeating the question on: “what would she use as the average shoe size?” Again, Lina mentioned that it would be five and justified her decision based on the “majority” bought this particular shoe size and because of the “average 5.4”. The above Excerpt L2 illustrates these points.

Table 4.5

Excerpt L3

R: What does average relate to in terms of the measures of central tendency?
S: Mode and mean
R: So, are you telling that the average can be associated to both the mode and mean?
S: Average is actually the mean
R: But earlier on, you mentioned that it can be also the mode
S: No no no… average is actually the mean. Mode is the most frequent.
R: But if I asked, what would you use as the average shoe size?
S: Size five
R: You base your answer on?
S: Most frequent and the formula
R: So, if the most frequent, don’t you think it is indicating the mode.
S: I am checking my answer. I think I screw up everything.
R: Don’t worry. Just explain to me your answer.
S: Average is actually the mean.
R: But when I asked you what you would use for average, you mentioned five and when I asked you what the average is, you gave me five point four. Can you explain this?
S: Yes. The average that I calculated is five point four. But in this case, the average, if it is not going to be five it is going to be four. We have to use a whole number because shoe size. I used the “most frequent” to decide on my average.

The above Excerpt L3 further illustrates Lina's response when the researcher probed her further with the question: what does average relate to in terms of measures of central tendency. Lina quickly answered mode and mean. However, when the researcher clarified if average can be referred to both the mode and the mean. Lina clearly distinguished that the average is actually the mean and the mode is "the most frequent".

The researcher delved further for clarification. Lina showed some confusion initially and mentioned that she needed to recheck her answers. However, she was assured not to worry but to just explain her answers. After which Lina firmly mentioned that the average is actually the mean.

Nevertheless, Lina had given two different answers to the questions on "what she would use as the average" in which she mentioned five and for the question on "what is the average", she mentioned "five point four". Finally, Lina explained clearly that her decision to use "five" as the average shoe size was because she knew that the shoe sizes had to be presented as whole numbers. Therefore, Lina utilised the "most frequent" idea to decide on the suitable whole number for the average shoe size from the calculated mean value.

Task 3 (Mode). Task 3 is another task that is used to explore the subject's knowledge of the mode that can represent a typical behaviour, a characteristic that can be used as a quick method of reporting central tendency or average. This task is presented in a tabular form.

Table 4.6

Task 3

The table below shows the number of comments made by eight students during a class period on a particular day.

Student	Number of comments
A	0
B	5
C	2
D	22
E	3
F	2
G	1
H	2

What is the mode?
How did you obtain this value as the mode?
If you were to represent one value for the number of comments made by the students on that day, would you use this value?
Why?

Lina had two different ways of identifying the mode. First, she identified the mode based on the "highest" number of comments but later she changed to the "most frequent" number of comments. Lina had some knowledge regarding the mode as a form of data representative.

Table 4.7

Excerpt L4

R: What is the mode?
S: It is student D.
R: How did you obtain this value as the mode?
S: Based on the highest number of comments
R: If I change this value, 22 to 6. What is the mode now?
S: Still the same which is student D.
R: How did you obtain the mode?
S: Based on the highest number of comments made by the students. Oh sorry [she looked uncertain now]. I want to change my answer.
R: What is your answer?
S: The mode is 2 because there are three students who made 2 comments.
R: Can you explain on how you obtain the mode?
S: Because by looking at the number of comments made by the students, the most number of comments made is "2". So, the mode is "2"
.
.
.
R: If you were to represent one value for the number of comments made by the students on that day, would you use this value?
S: No
R: Why?
S: Because if we look at the value, the number of comments made by students carry a quite wide difference. The lowest number of comments is zero and the highest is 6. So, the range is 6. Looking

at in terms of the standard deviation and also the dispersion, the standard deviation is roughly 3 point something. [Looked rather unsure of her answer] Can I think for a while?
R: Yes. You take your time.
S: It can be used. (Refers to the mode)
R: Why?
S: Because the difference between the numbers of comments is not that extreme (for the case where 6 is used) compared to the previous one which is 22.
R: What if for the case of 22?
S: In this case, I would not use this value because the difference is very extreme. One is zero and the other 22.

Lina initially identified the mode in the Task 3 as student D. She based her decision on the "highest" number of comments. When the researcher, changed the value "22" in the table to "6" and repeated the question. Lina mentioned that the mode is still student D in which she also based on the "highest" number of comments. However, Lina showed uncertainty towards her answers and requested to change her answers. Lina changed the mode to "2". She claimed now that she based her answer on three students who made "2" comments. Lina added that the "most" number of comments made is "2", therefore the mode is "2"

However, Lina mentioned that she would not use "2" (the mode) to represent the number of comments made by the students on that day. She explained the number of comments made by the students carried quite a wide difference with lowest number of comments as "0" and the highest number of comments "6". She added that the range is six and looking at both the standard deviation and dispersion, the standard deviation is roughly three point something. However, Lina looked rather unsure and she requested for some time to think.

After thinking about it for a while, Lina decided that the mode "2" can be used to represent the number of comments made by the students on that day. According to her this is because the difference between the mode "2" and the other number of comments is not extreme (for the case where the number of comments for student D is 6) as compared to the previous one which is 22.

However, Lina said that she would not use the mode “2” to represent the number of comments for the case where the number of comments for student D is “22”. She explained this is because of the extreme difference between the mode and two other data points which are “0” and “22”. The above Excerpt L4 describes this.

Task 4 (Average). Task 4 is used to explore the subject’s knowledge of the average. The average in the given context is based on the median. Here, the subjects’ knowledge of the extreme values and the robustness of the median over the mean in representing data involving a large population is explored.

Table 4.8

Task 4

A research study found that “Malaysian primary school students watch an average of 3 hours of TV per day”. However, the average was obtained after considering the information that a small group of students watched more than 12 hours of TV per day.

What does “average” mean in this sentence?
How do you think they got this average (3 hours of TV per day)?
Can you explain using an example?

Lina mentioned that the average can be based on the mean or the mode but not the median. She was aware of outliers and knew that the outliers can affect the reporting of an average.

Table 4.9

Excerpt L5

R: What does average mean in this sentence?
S: Almost all students watched 3 hours of TV per day. Plus, minus.
R: How do you think they got this average?
S: Taking into account, number of hours the students watch TV.
R: Can you explain using an example?
S: Student A 3 hours, B 4 hours and C 2 hours. Using the formula, you get 3 hours of TV per day.
R: How do you think the information that a group of students watched more than 12 hours of TV per day will influence the average?
S: Umm… (Looked unsure) It is going to increase the average.
.
.
.
R: What if I ask you this way? How do you think they got the average 3 hours of TV per day?
S: Using either mean or mode.
R: Are you saying that average can be based on mode?
S: It depends on the data. I think.

R: But then the data in this particular survey done also contain information saying that a small group of students watched more than 12 hours of TV per day. How was this information used here?
S: If I were to relate with what I learned, it is going to be what we call the outliers that we can ignore this data.
R: Can you explain further?
S: Ahhh…. because if I am not mistaken, the data contain outliers which is out of data. Let's say in 100 almost about 90 plus 1 2 3 4 hours and only 1 or 2 is more than 12 hours then we can just ignore the data because if we are going to use than it is going to affect the reporting of the data.
R: How does it affect the reporting of the data?
S: The average will be affected.
R: If we look at measures of central tendency, which measure would have been a better measure to be used in reporting this data?
S: Mode, I think.
R: Why mode?
S: Because majority of students watch almost 3 hours. So, if were to report, I mean to represent the whole population of the data, mode will be the best instead of average. Because it is stated that a group of students watched more than 12 hours of TV per day so if we were going to use average than it is going to affect the reporting of the value.
R: But earlier, you mentioned that the outliers can be ignored. What do you mean by this?
S: If they take into account 12 hours, it is going to affect the reporting of the average.
R: So, what is the best measure to be used for the reporting of the data?
S: Mode [thinks for a while]. Maybe median also can [unsure whether can or cannot] No. It is mode I think.

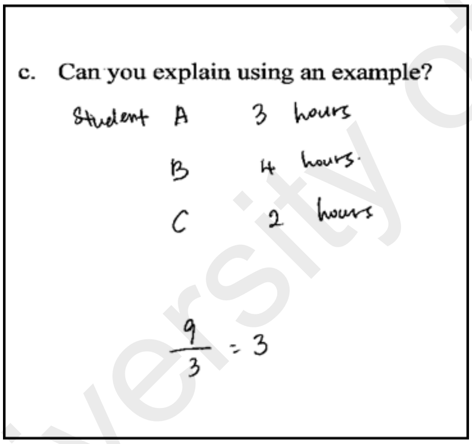


Figure 4.3. Entry L3

When Lina was asked to explain the meaning of the average in the above context, she mentioned as "almost all" and also "plus minus". Lina described the use of the arithmetic mean formula when she was asked on how the average would have been obtained in which she also wrote down in the task sheet shown in Entry L3.

However, Lina looked unsure at first when she was asked on how the information on a group of students watched more than 12 hours of TV per day will influence the average but later she mentioned that this information can increase the average. Lina mentioned that the average can be obtained by using either the mean or the mode. She claimed that the average can be based on the mode but it depends on the data. However, her answer here indicated some uncertainty when she mentioned "I think"

After Lina was probed deeper on the information of 12 hours of TV per day, she mentioned that this information contains outliers that can be ignored. She explained that if this particular data is not ignored than it will affect the reporting of the data. She further explained that what she meant by "affect the reporting of the data" was that the average will be affected.

Lina mentioned that among all three measures of central tendency, the mode is a better measure to report the data in this situation. She elaborated that the mode is more suitable because majority of the students watched almost 3 hours. Therefore, the mode is the best measure to represent the data instead of the average. According to her also, the information that a group of students who watched more than 12 hours of TV per day will affect the average if the average is used instead.

When the researcher asked Lina for clarification about her statement that "the outliers can be ignored". Lina mentioned that if this data is taken into account than it will affect the reporting of the average. Lina had some doubts on whether to consider the median as the measure for the reporting of the data. However, later Lina decided that the median is not suitable and that the mode is still the best measure. The above Excerpt L5 describes this:

Task 5 (Median). Task 5 is straightforward task that is used to explore the subject's knowledge of the median given in daily life context. This task is used to discuss the robustness of the median over the mean in the presence of an outlier.

Table 4.10

Task 5

The weights in kilograms for 9 children are 15, 25, 17, 19, 16, 26, 18, 19, and 24. Which is the weight for the median child? Which is the median if we include another child who weighs 43 Kg? Is it adequate to use mean to represent the weight of the 10 children? Why?

Lina knew that the median is a better measure compared to the mean to represent a set of data in the presence of an outlier. She knew how to handle an outlier if the mean is considered. This is found to be different from what Lina described in the earlier task.

Table 4.11

Excerpt L6

R: Which is the weight of median child?
S: 19
R: How did you get this value?
S: Median is something in the middle. We rearrange the data in ascending or descending order. Then we, divide the two middle values.
R: How about in this task?
S: So, the right and left, will have the same number of data. It is going to be 19 and 19, so the middle is 19.
R: Without including 43 kg. How did you get 19 there?
S: Sum it and divide by two. The average of two. (Points out to the average of the two middle values)
.
.
.
R: If we have included 43 kg. How?
S: 19. So here is four and here is four so 19. (Points to the middle value)
R: Is it adequate to use mean to represent the weight of the 10 children?
S: No.
R: Why?
S: Because the tenth child's weight is 43 kg and the difference between the data is too big. So it is not proper to use the mean.
R: What measure do you think will be suitable to be used to represent the weight of all 10 children?
S: Median.
R: Why median?
S: [pause a while] because the difference between the median and the rest of the data is not much.
R: Anything else that you would like to add on?
S: Or we just discard 43 because it is an outlier.
R: How you do this?
S: Yes, I think we can use mean. If we discard 43, we can use mean.
R: If we don't discard 43?
S: Median would be better.

Lina mentioned that the weight of the median child is 19. She said that the median is the middle value. After when, she elaborated that the data should be first arranged in the ascending or descending order. Then, the two middle values are divided.

Lina explained further after she was probed that equal number of data should be obtained both on the right hand and the left-hand sides of the arrangement. Later, she pointed at the two middle values which were both 19 and she mentioned that the middle is 19. She further explained that to sum up these two middle values and then divide the total by two. Lina also stated "the average of two".

For the case where 43 kg is included, Lina mentioned that the median is also 19. She explained that there are four data values on both sides of the ordered arrangement. She pointed out to 19 as the median, the middle value in the ordered arrangement. Here she did not perform any calculation or averaging.

Lina expressed that the mean is not adequate to represent the weight of all the 10 children because of the weight of the tenth child 43 kg. She elaborated that there exists a big difference between 43 kg with the other data in the set. Therefore, the mean is not proper to be used to represent the weight of all ten children.

However, Lina mentioned the median is a more suitable measure to represent the weight of all ten children. She explained that the median is a better measure because of the difference between the median and the rest of the data is not much. On the other hand, Lina mentioned that 43 kg is an outlier that can be discarded. She explained that the mean can be considered as a measure to represent the data if 43 kg is discarded. However, she also mentioned that if 43 kg is not discarded than the median would be a better measure. The above Excerpt L6 illustrates Lina's points on this.

Task 6 and task 7 (Idea of average). Task 6 and Task 7 are data construction tasks based on the idea of average. These tasks provide insights on the subjects' knowledge of the average and how they construct data sets based on the given average. These tasks are used to explore the subjects' representativeness of average whether as a typical value, a middle value, a balance point, a fair share or an algorithm.

Table 4.12

Task 6

The average price of a bag of crisps is RM 27. We have seven bags of crisps each of which has an empty price tag. Place a price on each of the bags so that the average price is RM 27.

Table 4.13

Task 7

The average price of a bag of crisps is RM 27. We have seven bags of crisps each of which has an empty price tag. Place a price on each of the bags so that the average price is RM 27. However, none of the bags of chips can cost RM 27.

Lina showed some knowledge of the average as a typical value in Task 6. In both the tasks Lina utilised the "backward" direction of the mean calculation to find the sum of all values from the given average. She utilised the sum of all values to guide her construction of data sets especially in Task 7.

Table 4.14

Excerpt L7

R: How did you place the price tag?
S: All RM 27.
R: Why did you do that? Can you explain?
S: The average is RM 27. So, it is actually plus minus the price is less than 27 or more than 27. It will easier if we put all as RM 27.
R: Any other reasons for placing all RM 27?
S: If looking at the aspect of the formula then it is going to be RM 27 times 7 divide by 7{} so everything is going to be 27. So, it is easier for me to use RM 27 instead of calculating one by one.
.

.
.
R: Now for Task 7. How did you get these values here?
S: Since the average is 27 so the price is around this value. It can be lesser such as 23 or 30 but still around 27. Plus, minus 27 that will sum up to 189.
R: How did you get each of these values?
S: I just put randomly any value.
R: Did you base these values on anything?
S: It is based on 27. Less than 27 or more than 27 but not too less or more because it will affect and can be outliers.
R: How did you get these values?
S: I just sum everything. Put any numbers that will add up to 189.
R: Any particular ways to determine these numbers?
S: No particular ways. Just randomly place as long as it adds up to 189.

For Task 6, when Lina was asked on how she placed the price for all seven bags of crisps, she mentioned "All RM 27". Lina elaborated that since the average price is RM 27, so the price of each bag has to be plus minus the average price, it can be either less than RM 27 or more than RM 27. Lina felt that placing the price of all the bags as RM 27 will just make things easier for her. She further elaborated that using the mean formula, 27 when multiplied by 7 and then divided by 7 will give her 27. So, she felt it is easier for her to use RM 27 instead of calculating one by one.

On the other hand, in Task 7, where a restriction was posed on using back the average price, Lina placed the prices of each bag of crisp around the average price 27. She mentioned that the price can be lesser than RM 27 such as 23 or more say around 30 but the price must still be around the average price which is 27. She also mentioned that the prices should be plus minus 27 and all the prices will have to sum up to 189.

However, when she was probed on how she got each of the values that she had written in the boxes, Lina claimed that the prices were randomly placed but based on 27. She elaborated that it can be any values that are less or more than 27 but emphasised that not too less or too more than 27. She mentioned that prices which are too far less or more than 27 can affect the data because these values can be outliers. Lina claimed that she had put the prices randomly but she made sure that the sum of these random prices will add up to 189. The above Excerpt L7 describes Lina's decisions on the data construction for both tasks.

Summary.

Table 4.15

Subject matter knowledge with reference to context

Task	Statistical element	Conclusion
Task 1	Mean as average	Lina's knowledge of the mean is the average. She mentioned that the average is obtained by using the sum and division. Lina elaborated on the algorithm to find the mean and even showed an example on the mean calculation. Lina mentioned that the use of the mean algorithm can produce a decimal average which might not be logical for the context of the given data.
Task 2	Idea of mode	Lina related the idea of mode to real-life situation without the task explicitly stating so. She utilised the frequency tally to obtain the mode. She identified the mode based on the "most frequent" data.
	Mode as data representation	Lina said that the mode can be utilised as a form of data representation for the shoe size in her shop because "majority" bought that particular shoe size.
	Mode as average	Lina did not utilise the mode as a quick method to report an average. Instead, she only used the mode as a guide to decide on the average. Lina based the average on the mean. Lina calculated the average using the mean calculation. Lina realized that for the context of the data which was referring to shoe sizes, the ideal average has to be a whole number. Since the average obtained was in the form of decimal number, she had to decide on the suitable whole number for this average. Therefore, the mode was used to guide her decision on the suitable whole number.
Task 3	Idea of mode	Lina initially identified the mode based on the highest number of comments but later she changed to the "most frequent".
	Mode as data representation	Lina said that she would only utilise the mode as a form of data representation for situations where all the other individual data values are close to the mode value. In the case where there exist data points which are far from the mode, Lina mentioned that the mode would not be a suitable form of data representation.
Task 4	Median as average	The reporting of an average on data related to human population is normally based on the median. However, in this task, Lina mentioned that the average can be based on the mode or the mean. She did not mention that the average can be based on the median. Lina said that the average can be based on the mode but it depends on the data. Lina's responses to the later questions indicated that her knowledge of the average is synonymous to the

		mean. She claimed that the mode can be a better measure to report the data in Task 4 compared to the average because of the influence of outliers on the average which indicated that the average she was referring to is the mean.
Task 5	Idea of median	Lina elaborated the median procedure both for even and odd numbered data. She said that the data is arranged in ascending or descending order first and then if the data is odd numbered, the median would be the middle number in the arrangement. Otherwise if the data is even numbered, the median is the mean of the middle two number in the arrangement.
	Robustness of measures	Lina mentioned that in the presence of an outlier in a set of data, the mean is not adequate to represent this data. As an alternative, she mentioned that the median is more appropriate because the difference between the median value and the rest of the data is not much. Furthermore, Lina explained that the mean can be considered if the outlier is discarded but if the outlier is not discarded than the median would be a better measure to represent the weight of all ten children.
Task 6	Idea of average	In the case, where the average value is not restricted from being used as a value in the data set, Lina showed that the average as some sort of a typical value. Lina utilised the average value as all the values in the data set. She based this on the "backward" mean calculation where she found the total from the given average and then divided the total again with the number of values needed. Therefore, she obtained again the mean value which she placed as the values in the data set.
Task 7	Idea of average	On the other hand, when a restriction was posed that the average value cannot be repeated as a value in the data set, Lina used the total value as a guide to ensure that the sum of all individual values that she placed is within the total. She utilised the backward mean calculation to get the total. Here, Lina placed the values randomly but bearing in mind that the values should be around the average value. Not too far less or more because if not these values will be regarded as outliers.

Subject matter knowledge of measures of central tendency in handling bias. This construct is used to investigate the subject matter knowledge of the measures of central tendency involved in handling bias. One important link between the measures of central tendency and statistical literacy is the discussion on the extreme values and which of the three measures is most or least likely to be biased by these values. Knowing the distinctive features of the mean and the median can lead to an appreciation on why it may be important to use one or the other in the presence of outliers. A total of two task were constructed to explore this construct.

Task 8 and task 9. Task 8 and Task 9 are used explore the subject's knowledge of choosing the most suitable measure of central tendency to represent a set of data. These tasks are used to explore how the subject deals with the issue of bias involving the measures of central tendency in the presence of an outlier and whether the subject is able to use the data to give an indication of the most suitable representative measure. The presence of an outlier does have a biasing effect on the mean. In such situation, the mode or the median is least affected. In fact, the median is regarded as a more robust measure to represent a set of data that contains an outlier.

Table 4.16

Task 8

A small object was weighted on the same scale separately by nine students in a science class. The weights (in grams) recorded by each student are shown below.

6.3 6.0 6.0 15.3 6.1 6.3 6.2 6.15 6.3

The students want to determine as accurately as they can the actual weight of this object.
What are the measures of central tendency that you can obtain based on the above data?
Which is the best measure of central tendency to represent the actual weight of this object?
Why?

Table 4.17

Task 9

Five employees in a company were chosen randomly for a study about salaries. They obtained the following salaries:

Employee	A	B	C	D	E
Salary (per month)	54,000	42,000	60,000	48,000	180,000

What are the measures of central tendency that you can obtain based on the above data?
Which is the best measure of central tendency to represent the actual salary of the company? Why?

Lina identified the measures of central tendency that can be obtained from the given sets of data. She also chose on the best measure of central tendency to represent the data and justified her decision for choosing that measure.

Table 4.18

Excerpt L8

R: What are the measures of central tendency that you can obtain based on the above data?
S: Mean 7.18, Mode 6.3, Median 6.2
R: How did you obtain these values?
S: Mean, total up everything and divide with the number of data, median is the middle value after arranging in ascending order and mode is based on the highest frequency.
R: Among these measures of central tendency, which is the best measure of central tendency to represent the actual weight of this object.
S: Either mode or the median
R: Why?
S: Because the value represents the data. It is not proper to use the mean because we have 15.3, a bit away from the majority of the data.
R: How do you know this?
S: Because eight of the rest data except 15.3 are around the mode or median value. The values are plus minus 6. So, it is around this.

Lina mentioned that the mean, the mode and the median can be obtained based on the data give in Task 8. She elaborated on the mean calculation as total up all the given data and then divide the total with the number of data. She also elaborated the median is the middle value after the data is arranged in ascending order. She mentioned that the mode is based on the highest frequency.

Lina mentioned that either the mode or the median can best represent the actual weight of the object. She added that the mean is not suitable because of one particular data which is 15.3. She said that 15.3 is bit away from the majority of the data.

According to Lina, eight out of nine data except 15.3, are values around the mode or the median values which are plus minus 6. Therefore, the mode or the median are more suitable measures to represent the given data. The above Excerpt L8 illustrates Lina's point on this.

Table 4.19

Excerpt L9

R: What are the measures of central tendency that you can obtain based on the above data?
S: Mean 76 800, No mode and median is 54 000
R: How did you obtain these values?
S: Mean, total up everything and divide with the number of data, median is the middle value after arranging in ascending order and mode is based on the highest frequency.
R: But you mentioned no mode. Why no mode?
S: Because the frequency occurring each employee's salary is one. So, no highest frequency here, so no mode.
R: Which is the best measure of central tendency to represent the actual salary of the company?
S: The median
R: Why?
S: Ignoring the 180 000 then the rest of the salary is plus minus around this median value.
R: Please explain further.
S: Because the salary is around 50 000 and ignoring the 180 000.

Similarly, in Task 9, Lina also identified the measures of central tendency that can be obtained based on the given set of data. She elaborated on the mean calculation as total up all the given data and then divide the total with the number of data. She also elaborated the median is the middle value after the data is arranged in ascending order. She mentioned that the mode is based on the highest frequency.

Lina mentioned that there is no mode here because the frequency for each employee's salary is one. Therefore, she mentioned that there is no highest frequency here. According to Lina, the median is the best measure of central tendency to represent the data. She explained that except for 180 000, the rest of the salaries are plus minus around the median salary which is around 50 000. The above Excerpt L9 illustrates Lina's point on this.

Summary.

Table 4.20

Subject matter knowledge in handling bias

Task	Statistical element	Conclusion
Task 8	Identifies and summarises data using mean	Lina utilised the mean algorithm to calculate the mean.
	Identifies and summarises data using median	Lina carried out the odd median procedure to find the median
	Identifies and summarises data using mode	Lina found the mode based on the highest frequency.
	Best data representation measure	Lina mentioned that in the presence of an outlier, the mean is not suitable to represent the data. Instead, she mentioned that the median or the mode is more suitable. She said that the mode and the median values represent the data because these values are around the rest of the data except the outlier. Lina looked at all the individual values and decided on the best measure that can represent the data.
Task 9	Identifies and summarises data using mean	Lina utilised the mean algorithm to calculate the mean.
	Identifies and summarises data using median	Lina carried out the odd median procedure to find the median
	Identifies and summarises data using mode	Lina found the mode based on the highest frequency.
	Best data representation measure	Lina mentioned that in the presence of an outlier, the mean is not suitable to represent the data. Instead, she mentioned that the median is more suitable. She said that the median value represents the data because it is around the rest of the data except the outlier. Lina looked at all the individual values and decided on the best measure that can represent the data.

Subject matter knowledge of measures of central tendency in problem solving. This construct is used to explore the subject's knowledge of solving weighted mean problems. In order to solve the weighted mean problems, the subject's knowledge of applying the mean algorithm in both "forward" and "backward" direction is explored. Furthermore, knowing the representative nature of the mean in making sense of why data sets need to be combined in weighted mean problems is also explored. Although, algorithms are necessary but understanding the connections of the algorithms with the purposes are equally important too. These are considered as the more technical and sophisticated goals of statistical literacy. A total of two tasks were constructed to explore this construct.

Task 10 (Weighted Mean). Task 10 is a weighted mean problem which does not explicitly state to solve for the weighted mean. The task is used to explore the subject's ability to solve the weighted mean by combining the algorithmic numerical skills with the appreciation of the context within which they are used. Solving a weighted mean problem also involves making sense of why one would want to combine data sets to obtain the weighted mean.

Lina solved for the weighted mean without the task explicitly stating to find one. She knew the representative nature of the mean in making sense of why data sets need to be combined in the weighted mean problems.

Table 4.21

Task 10

A student attended College A for two semesters and earned a 3.2 GPA. The same student attended College B for three semesters and earned a 3.8 GPA. What was the students' overall GPA?

Table 4.22

Excerpt L10

R: What was the students' overall GPA?
S: 3.56
R: How did you obtain the overall GPA?
S: Umm sum the two and divide by 5. I mean for college A, the GPA is 3.2 so actually for two semesters she actually has 6.4. For another college, she 3 times 3.8 for three semesters which is 11.4. Add up the two divided by five semesters so you get 3.56.
R: Why did you times 3.2 with 2?
S: for the two semesters
R: Why did you times 3.8 with 3?
S: that's for three semesters
R: And why you divide by five?
S: because the total semesters involved is five.

Lina mentioned that the overall GPA is 3.56. She mentioned "sum the two and divide by five" when she was asked on how the overall GPA was obtained. Lina elaborated that for College A, the given GPA is 3.2 for two semesters so the total grade point is 6.4 which she obtained from 3.2 times with 2. Similarly, she mentioned that for College B the given GPA is 3.8 for three semesters and the total grade point is 11.4 which she obtained from 3.8 times with 3. She said that she added up the two total grade point for both the colleges and then divided the overall total with five semesters to get the overall GPA as 3.56. The above Excerpt L10 describes Lina's point on this.

Task 11. Task 11 is also a weighted mean problem similar to Task 10. The task is used to explore the subject's ability to solve the weighted mean by combining the algorithmic numerical skills with the appreciation of the context within which they are used. Solving a weighted mean problem also involves making sense of why one would want to combine data sets to obtain the weighted mean.

Table 4.23

Task 11

A recent study found that:
25 rural students watched an average of 8 hours of TV per weekend;
75 urban students watched an average of 4 hours of TV per weekend.
Show clearly the average TV viewing time per weekend for all 100 students.

Lina solved for the weighted mean despite the task asking to find the average. She also knew the representative nature of the mean in making sense of why data sets need to be combined in the weighted mean.

Table 4.24

Excerpt L11

R: How did you get this value?
S: Total, the 25 students with the average 8 hours multiply this and do the same for urban student then divide by 100
R: Why are you multiplying 25 with 8?
S: Because the average 8 is for 25 rural students and the same for urban, the average 4 is for 75 students.
R: Why are you adding these two values?
S: Because the question asks for 100 students
R: What does these values give you?
S: Sum of hours students watch TV per weekend for all 100 students.
R: Why are you dividing by 100?
S: Because I want to find the average for 100 students

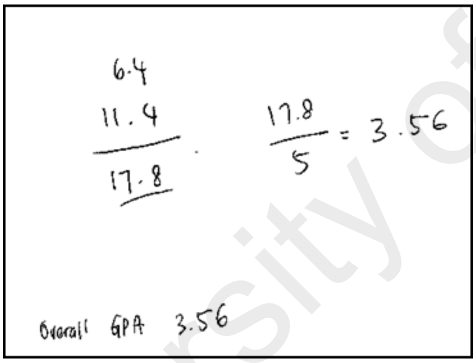


Figure 4.4. Entry L4

Lina wrote clearly on the calculation that she used to find the average as shown in Entry L4. When Lina was probed on how she obtained the average as 5 (by pointing at Lina's written answer in the task sheet), she explained that she totalled the multiplication of 25 with 8 for rural students and 75 with 4 for urban students and then divided the total with 100.

Lina explained that the given average 8 is for 25 rural students and the average 4 is for 75 urban students. She elaborated that she added the two totals that she obtained after multiplying the average with the number of students both for rural and urban because the question is asking for the average for all 100 students. Lina also mentioned that this gave her the sum of hours students watch TV per weekend for all 100 students. After when Lina said that she divided the sum with 100 because she wanted to get the average for all 100 students. The above Excerpt L11 illustrates Lina's points on this.

Summary.

Table 4.25

Subject matter knowledge in problem solving

Task	Statistical element	Conclusion
Task 10	Backward mean calculation	Lina solved for the weighted mean without the task explicitly stating to do so. She carried out the "backward" mean calculation, where she obtained the total for each group from the given group means.
	Representative nature of the mean	Lina used the knowledge related to the representative nature of the mean in making sense of why data sets need to be combined in the weighted mean.
	Forward mean calculation	Lina also carried out the "forward" mean calculation to calculate the weighted mean.
Task 11	Backward mean calculation	Lina solved for the weighted mean without the task explicitly stating to do so. She carried out the "backward" mean calculation, where she obtained the total for each group from the given group means.
	Representative nature of the mean	Lina used the knowledge related to the representative nature of the mean in making sense of why data sets need to be combined in the weighted mean.
	Forward mean calculation	Lina also carried out the "forward" mean calculation to calculate the weighted mean.

Subject matter knowledge of measures of central tendency in making inference. Measures of central tendency used to reduce or summarise data to single value(s) leads to the inference stage of any statistical investigation. Therefore, this construct explores the subject's knowledge on this. The construct is also used to explore the subject's knowledge of utilising the most suitable measure of central tendency as a significant tool for data comparison and draw inference based on the data without being specifically told.

Among all the three measures of central tendency, the mean is the most suitable measure that can be used to compare and draw conclusion about two data sets regardless equal or unequal sized especially when involving the performance of a player or a class. The mean has an advantage because it takes into account every data value in the data set. Therefore, the mean represents all the data in the data set and this characteristic of the mean can lead to a better conclusion especially of those related to the performance of players or class.

Task 12 (Equal-sized data). One of the goals of statistical literacy is to have a range of statistical tools in which the measures of central tendency can be used in summarising, comparing and drawing conclusion based on data sets without the tasks explicitly stating to do so. Task 12 is used to explore how the subject compared two equal sized data sets presented in numerical form. The task is used to explore on the subject's choice of a measure that is used to compare and draw conclusion about the data sets involving the performance of the players.

Table 4.26

Task 12

Coach Ahmad is selecting students to play on the All Star Team. He has decided to look at the scoring of each player during the last three weeks of the season. Below are the points scored by Ramu and Beng. If Coach Ahmad can only select one of the two players, who would you recommend he select and why?

Ramu:	21	16	23	21	20	17
Beng:	24	18	21	18	22	28

Lina utilised all three measures of central tendency to summarise the above data without the task stating to do so. She compared the consistency of these measures and also the scores to decide on the player of her choice.

Table 4.27

Excerpt L12

R: Who would you recommend him select?
S: Ramu
R: Why Ramu?
S: because his performance is quite consistent.
R: Why are you saying that his performance is quite consistent?
S: Based on the measures of central tendency, mean, mode and median.
R: Can you explain further?
S: His mean is 19 or 20, mode is 21 and median is 20 or 21. They are around the same values.
R: What about Beng?
S: Beng, the mean and the median is almost same around 21/22, only that the mode is 18. So even though Ramu's total score is less compared to Beng's total score but the consistency of him making the individual scoring, I mean the mode is higher compared to Beng.
R: So, whom do you choose?
S: Ramu
R: Why Ramu?
S: His performance is consistent
R: What do you mean by performance consistent?
S: His mean, mode and median are around the same value
R: So How did you choose Ramu?
S: Based on all three values actually.
.
.
.
R: If you had to base on one measure?
S: I would use the mode
R: Why?
S: Because it gives the one's highest frequency score.
R: So, who will you chose based on the mode?
S: Ramu
R: Why?
S: Because his mode is 21 compared to Beng which is 18
R: What if in this data, there is no mode existing. (The researcher modifies the task so that the data has no mode)
S: Then, I will go for mean or median. But I think I will go for the mean.
R: Why?

S: Because that will be his average score. So, the possibility of him scoring plus minus the mean score is high I think

Lina recommended Ramu to be selected for the team. According to her, his performance was found to be quite consistent. She said that the consistency of his performance was based on all three measures of central tendency, the mean, the mode and the median. She mentioned that Ramu's mean was 19 or 20, his mode was 21 and his median was 20 or 21. She found that these values were around the same values.

Lina elaborated that for Beng, his mean and median were almost same around 21 to 22 but his mode was only 18. She added that even though Ramu's total score was less than Beng's total score but his consistency of scoring the individual game which she meant the mode is higher compared to Beng.

Therefore, Lina chose Ramu based on his consistent performance. She repeated again that what she meant as consistent performance was that his mean, mode and median values were found to be around the same values. She had actually utilised all three measures of central tendency to choose Ramu.

However, when Lina was asked to choose one particular measure and decide on the player. Lina chose the mode. She explained that it would give one's highest frequency score. Again, she mentioned that Ramu would be chosen because he had higher mode than Beng. The researcher delved further by changing the situation to a "no mode situation". Lina was asked on how she would decide now. First Lina mentioned that she would use either the mean or the median but later she decided that she would use the mean. She said that the mean would give his average score and the chances of him scoring the next game will be plus minus the mean score is high. However, Lina's response indicated uncertainty when she mentioned "I think". The above Excerpt L12 describes this.

Task 13 (Unequal-sized data). Task 13 is used to explore how the subject compared two unequal sized data sets presented in numerical form. The task is used to explore on the subject's choice of a measure that is used to compare and draw conclusion about unequal sized data sets involving the performance of the players.

Table 4.28

Task 13

Coach Ahmad would like to select another player. Who should he select from the group below and why?

Ali:	25	19	28	25	23	16	18	24
Khan:	26	16	27	16	29	24		

Lina looked at the scores to select the player for the team. She compared the individual scores of the players and also the consistency of scoring by the players.

Table 4.29

Excerpt L13

R: Look at task 13. If Coach Ahmad has to select another player. Who would he select from this group now?
S: I would choose Ali.
R: Why?
S: Because of the consistency of his performance. If we look at Khan's performance, the gap between the score is quite a bit. It is going to be a bit risky if he were to choose Khan.
R: How did you base your judgement?
S: Logic. [Laughed]
R: I see that you have looked at the three measures of central tendency.
S: Yes. I depend on the data.
R: How?
S: I did show the measures of central tendency, it is just that when I looked back the data, the difference between the score for Khan, the gap is quite big so I did not use the mean, mode or the median.
R: How did you base then?
S: I used the scores and I chose Ali because his scores are quite consistent.
R: If you were to use the measures of central tendency. How?
S: I will choose the mode.
R: Why?
S: The mode for Ali is 25 and Khan is 16. I did not use the mean in this case because of the spread of Khan's data. Even though when we look at the mean, we will surely choose Khan but because of the spread of data that's why I ignored the mean and I chose the mode. Looking at the mode, the score for, the frequent score for Ali is 25 and Khan is 16. That is why I chose Ali.
.
.
.
R: So, if the same situation exists in Ali and Khan's data. No mode exists. (The researcher modifies the task so that no mode exists). How do you select the player then?

S: I will use either mean or median. But I have to look through the data, the spread.
R: Which is the best?
S: It is the mean I think.
R: Why?
S: Because plus minus that is the score he gets.
R: Can you explain further?
S: The player's score will also be around his average score. So the tendency of his next score can be based on his average score.

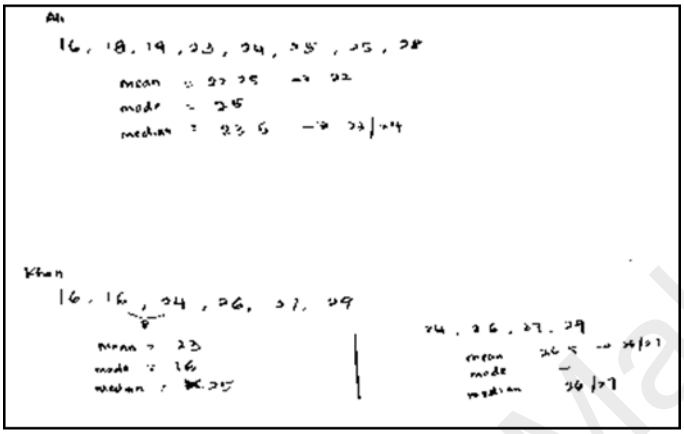


Figure 4.5. Entry L5

Lina chose Ali for the team. Lina said that Ali was consistent in his performance whereas Khan had quite a bit of gap in his scores. Therefore, Lina felt that Khan is a bit risky to be in the team. She mentioned that she based her judgement on logic but she laughed at her response.

Lina had found the values for all three measures of central tendency in which Lina had written in the task sheet as shown in Entry L5. However, Lina responded that she had actually looked at the data to decide. She said that although she calculated the measures of central tendency but when she looked at Khan's data, she noticed that the differences in his scores were quite big. Therefore, she said that she did not use the measures to decide. She mentioned that she used the scores and chose Ali because his scores were quite consistent.

However, when Lina was asked to decide based on the measures of central tendency, Lina chose to use the mode. She explained that she would not use the mean because of the spread in Khan's data. She explained that if the mean was used than Khan will surely be chosen. However, due to the spread of Khan's data, the mean was ignored and the mode was considered. Therefore, she based her decision by looking at the mode. Ali's mode was 25 meanwhile Khan's mode was 16. Based on this Lina, chose Ali.

Lina was probed further with "no mode" case. Here, Lina first mentioned that she would either use the mean or the median. But she also mentioned that she will look at the spread of the data. However, she mentioned that the mean is the best because his next score will be plus minus his mean score. She elaborated that the player's next score will be around his average score and the tendency of his next score can be based on his average score. The above Excerpt L13 illustrates Lina's points on this.

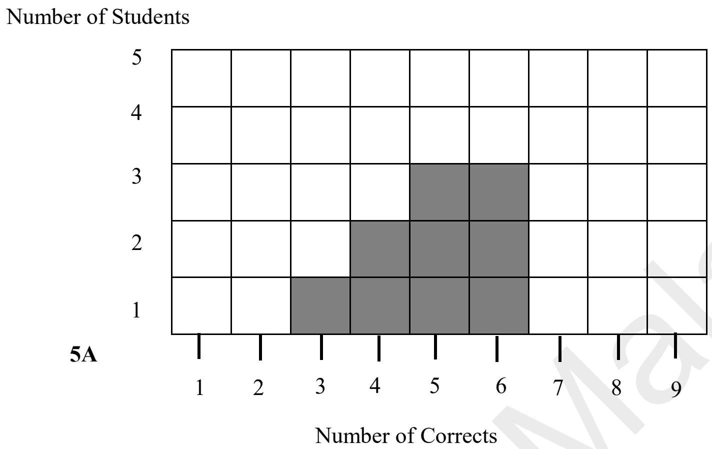
Task 14 (Equal sized data). "Initial experiences with inference can be linked with predictions from data sets presented in graphical form…" (Watson, 2006, p. 190). Task 14 is used to explore how the subject compared two equal sized data sets presented in graphical form without the task explicitly stating the comparison tool to be used.

Table 4.30

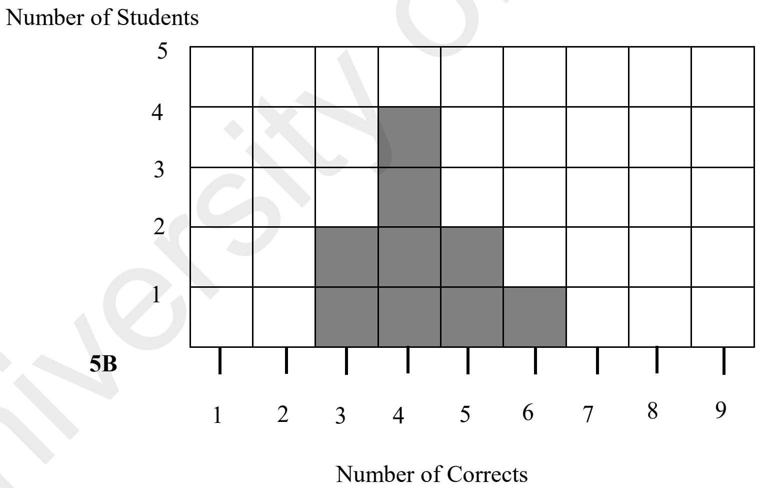
Task 14

A mathematics quiz was given to all Year 5 students of Sekolah M. Their test scores based on the number of corrects obtained were recorded. Random data sets of the test scores of two groups of students from class 5A and 5B were taken. The graphs below show the test scores for both the groups. Based on the graphs below, which class did better? Why?

a.



b.



Lina had utilised the measures of central tendency to compare two equal sized data presented in graphically form without the task explicitly stating the comparison tool.

Table 4.31

Excerpt L14

R: Based on the graphs, which class did better? Why?
S: 5A did better because if we look at in terms of the mean for the two groups, the mean for 5A is higher compared to 5B and so as for the mode and the median. The mode and median for 5A is more compared to 5B. So, it is 5A.
R: You mentioned here mean, mode and median. How did you base the performance of the class?
S: For this case can use all the mean, mode and median.
R: You said use all. But if you were to use one particular measure, how?
S: The mean
R: Why?
S: Usually we use mean
R: Why is that usually mean is used?
S: Mmm…. [Looked unsure] I change to mode
R: Why?
S: Because mode shows the highest number of students [paused a while and was thinking]. Mode shows the highest frequency of numbers of students answering the questions correctly. So, if more students can answer meaning is better is it [Sounded unsure]
R: Anything else?
S: Actually, we can use all. The mean, mode and median
R: But compared to all these measures, which one is the best to be used?
S: The mode
R: Why?
S: As I mentioned it shows the highest frequency of numbers of students answering the questions correctly. So, if more students is better.
.
.
.
R: Is there any other ways to determine the class that did better?
S: No. Only that

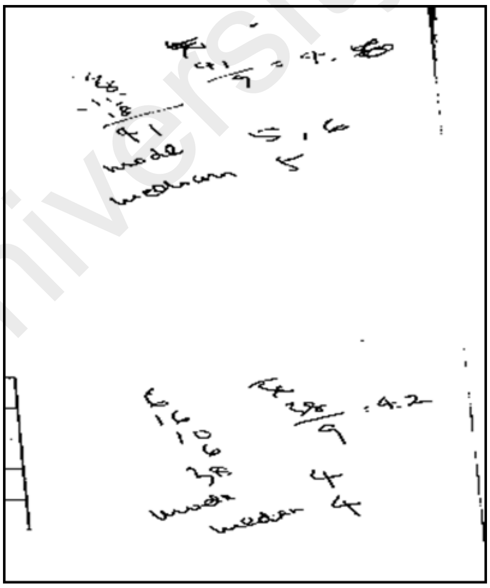


Figure 4.6. Entry L6

Based on the graphs presented in Task 14, Lina mentioned that 5A performed better. She compared the mean of the two classes and mentioned that the mean for 5A is higher compared to the mean for 5B. She also mentioned that the mode and the median for 5A are also better than 5B.

Lina mentioned that she based her decision on all three measures of central tendency. However, when she was asked to decide on one particular measure, Lina mentioned the mean. She explained that usually the mean is considered. When she was asked to explain further, she looked rather unsure of her answer and quickly changed her answer to the mode.

She mentioned that the mode showed the highest frequency of number of students answering the questions correctly. Therefore, she believed if more students can answer then it is better. But her response showed that she was unsure about this. However, Lina kept insisting that all the three measures can be used to decide but if she were to choose one particular measure than it would be the mode. Lina mentioned that she is not aware of any other ways to compare the graphs and draw conclusions. The above Excerpt L14 and Entry L6 shows Lina's points on this.

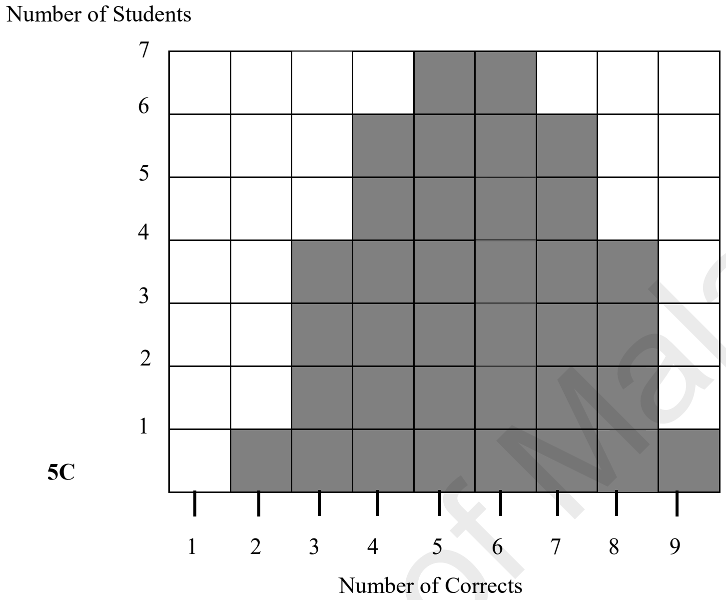
Task 15 (Unequal sized data). "Initial experiences with inference can be linked with predictions from data sets presented in graphical form…" (Watson, 2006, p. 190). Task 15 is used to explore how the subject compared two unequal sized data sets presented in graphical form without the task explicitly stating the comparison tool to be used.

Table 4.32

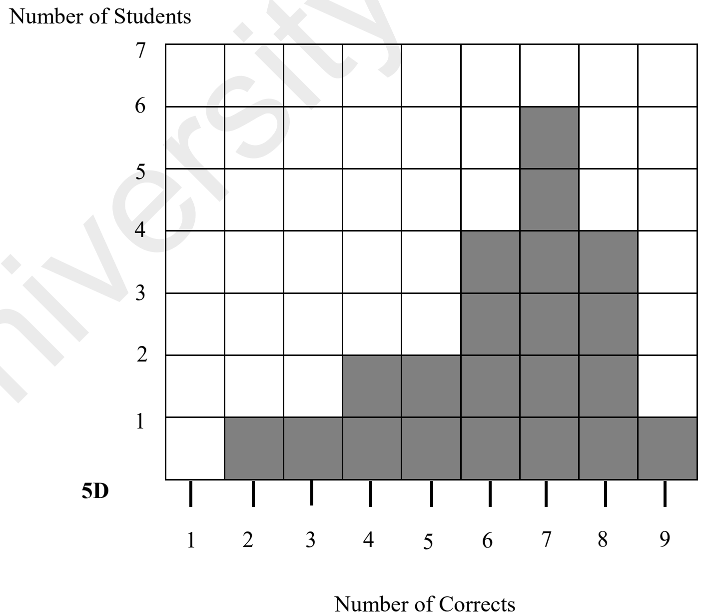
Task 15

Another two random data sets of the test scores of two groups of students from class 5C and 5D were also taken. The graphs below show the test scores for both these groups. Based on the graphs below, which class did better? Why?

a.



b.



Lina utilised the mode to compare two unequal sized data presented in graphical form. In fact, Lina believed that the mode is a better measure to compare not only two unequal sized data but also for cases where the data is skewed.

Table 4.33

Excerpt L15

R: Based on the graphs, which class did better? Why? S: 5D. Because in terms of the mode, the mode shows that 5D students answered more corrects as compared to 5C R: Any other way to determine which class did better? S: I would go for mode R: Why? S: Because more students in 5D answered more correct questions as compared to 5C. If look at the histogram, the 5C data looks normal but the 5D data looks more skewed. So, I don't prefer to use mean or the median. R: Why? S: Because when the data is skewed so the spread of the data is uneven so there can be outliers which can affect the calculation of the mean. It can also affect the median

Lina chose 5D as the class that performed better. Her decision was based on the mode. Lina mentioned that the mode showed that 5D students answered more questions correct as compared to 5C. When she was asked if there were other ways to determine which class did better, Lina insisted that she would still use the mode. She elaborated that based on the graphs, 5C data looked normal but 5D data looked skewed. Therefore, she did not prefer to use the mean or the median. She explained that for the data which is skewed, the spread of the data is uneven. The data can contain outliers which can affect the calculation of the mean and also the median. The above Excerpt L15 illustrates Lina's points.

Summary.

Table 4.34

Subject matter knowledge in making inference

Task	Statistical element	Conclusion
Task 12	Summarises equal sized numerical data using measures of central tendency	Lina used all three measures of central tendency to summarise equal sized numerical data without the tasks explicitly stating to do so.
	Utilises the appropriate measure to compare and draw conclusion about equal sized numerical data	Lina initially utilised all three measures to compare and draw conclusion based on the equal sized numerical data. Lina mentioned that if she had to choose one particular measure to decide then she would use the mode. She said that she would only use the mean in the case of no mode.

Task 13	Summarises unequal sized numerical data using measures of central tendency Utilises the appropriate measure to compare and draw conclusion about unequal sized numerical data	Lina used all three measures of central tendency to summarise unequal sized numerical data without the tasks explicitly stating to do so. Lina first compared the individual scores to decide on the player to be selected. When she was asked to decide on one particular measure, she chose the mode because she claimed that the mean is not suitable due to the spread of the data. Lina mentioned that only in the case of no mode then she would utilise the mean.
Task 14	Summarises equal sized graphical data using measures of central tendency Utilises the appropriate measure to compare and draw conclusion about equal sized graphical data	Lina used all three measures of central tendency to summarise equal sized graphical data without the tasks explicitly stating to do so. Lina compared all three measures of central tendency to decide on the performance of equal sized graphical data. When she was asked to choose one particular measure to decide, she chose the mean first because she claimed that usually the mean is used. Later Lina changed her mind and said that she would use the mode instead. She said the mode showed the highest frequency of the number of students with correct questions.
Task 15	Summarises unequal sized graphical data using measures of central tendency Utilises the appropriate measure to compare and draw conclusion about unequal sized graphical data	Lina used the mode to summarise unequal sized graphical data without the tasks explicitly stating to do so. Lina utilised the mode to compare and draw conclusions on two unequal sized graphical data. Lina mentioned that the mean or the median is not suitable to be used when the data is skewed because of the outlier.

Level of subject matter knowledge of measures of central tendency. In order to understand a pre-service mathematics teacher's level of subject matter knowledge of measures of central tendency for each of the four constructs, the coding process adapted from Learning Mathematics for Teaching (LMT) (2006) project by Hill, Ball, Bass and Schilling was used.

The following elaborates on the coding process. The researcher determined whether the statistical element was present (P) or not present (NP). If the statistical element was present (P), then mark: (a) appropriate (A) if the pre-service mathematics teacher's use of the statistical element was appropriate, accurate, or correct; or mark (b) inappropriate (I) if the pre-service mathematics teacher's use of the statistical element was inappropriate, inaccurate, or incorrect. If the statistical element was not present (NP), then mark (a) appropriate (A) if the absence of the statistical element seemed appropriate or not problematic; or mark (b) inappropriate (I) if the absence of the statistical element seemed inappropriate or problematic (i.e., the statistical element should have been present) (adapted from LMT, 2006). The following diagram illustrates the coding process thought.

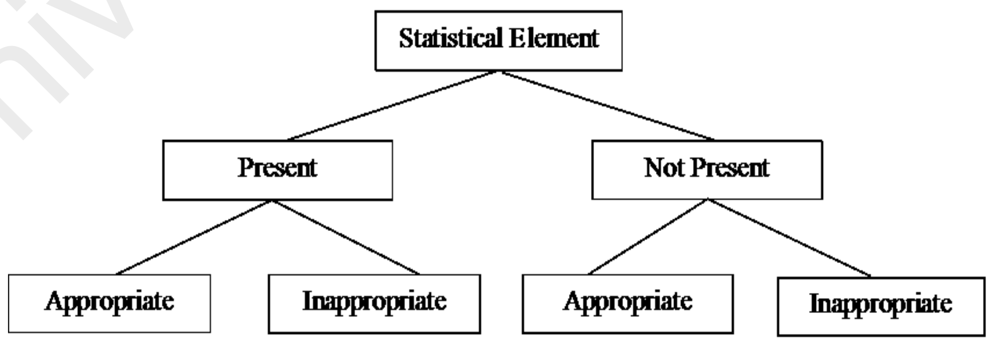


Figure 4.7. Coding thought process

A preservice mathematics teacher's level (low, medium, high) of subject matter knowledge of measures of central tendency for each of the construct was determined based on the percentage of appropriate statistical elements of the subject matter knowledge of measures of central tendency. The following illustrates the computation of the percentage of appropriate statistical elements of subject matter knowledge of measures of central tendency:

Percentage of appropriate statistical elements of subject matter knowledge of measures of central tendency was calculated based on the following formula = $\frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)}$ x 100%, where nPA, nPI, nNPA, and nNPI represents the number of codes that are coded as "present and appropriate (PA)", "present and inappropriate (PI)", "not present and appropriate (NPA)", and "not present and inappropriate (NPI)", respectively.

A preservice mathematics teacher's level (low, medium, high) of the overall subject matter knowledge of measures of central tendency was determined using the grading system used in the university where the data of the study was collected. In order to obtain a preservice mathematics teacher's level of subject matter knowledge of measures of central tendency, first the mean percentage of the subject matter knowledge for all four constructs namely; with reference to the context, in handling bias, in problem solving, and in making inference was calculated.

The following describes on how the mean percentage was calculated for all four constructs. Mean percentage of the subject matter knowledge of measures of central tendency = $\frac{A+B+C+D}{4}$ % where A represents the percentage of subject matter knowledge with reference to context, B represents the percentage of subject matter knowledge in handling bias, C represents the percentage of subject matter knowledge in problem solving and D represents the percentage of subject matter knowledge in making inference.

If a preservice mathematics teacher achieved 80% and above as the mean percentage of the overall subject matter knowledge he or she was assigned a high level of subject matter knowledge of measures of central tendency. If a preservice mathematics teacher's achievement was 40% to less than 80%, he or she was assigned a medium level of subject matter knowledge of measures of central tendency. Finally, if a preservice mathematics teacher's achievement was less than 40%, he or she was assigned a low level of subject matter knowledge of measures of central tendency.

Level of subject matter knowledge of measures of central tendency with reference to context. In Task 1, Lina elaborated on the arithmetic mean when she was asked on how the average would have been obtained. She also showed the mean calculation in her example. Lina knew that a decimal average is not logic given to the context of the data. However, she said that the decimal average can be produced as a result of the mean calculation without any consideration given to the context. Therefore, Lina had presented the knowledge of the mean as the average (P) which is considered as appropriate (A).

In Task 2, Lina had utilised the idea of mode without the task explicitly stating so. She carried out the frequency tally and had identified the mode as 5 based on the highest frequency. Therefore, Lina's idea of the mode was found present (P) and correct (A).

Lina said that the mode can be a form of data representation. She explained that this was because majority bought the mode size. Therefore, Lina's knowledge of the mode as a form of data representation was found to be present (P) and appropriate (A).

However, Lina utilised the mode as a guide to decide on the average and not as a quick method in the reporting of an average. In fact, Lina calculated the mean. Since the mean value that she obtained was in decimal form, Lina had to decide on a suitable whole number to round off the decimal average. She knew that for the context of the data which referred to shoe sizes, a decimal average was not suitable. So, Lina utilised the mode to guide her decision on the suitable whole number for the average shoe size. Therefore, Lina's knowledge of the mode as the average was not only found to be not present (NP) but also inappropriate (I).

In Task 3, Lina initially identified the mode based on the "highest number of comments" but later she corrected her response by identifying the mode based on the "most frequent" comments. Lina said that the mode is now 2 because the most number of comments made is 2. Therefore, Lina's idea of mode was found to be present (P) and correct (A).

Lina said that the mode can be used to represent a set of data. Nevertheless, she justified that this idea is only applicable if the mode and all the other values in the data are close to one another. Lina mentioned that if the number of comments for student D was 6 then the mode "2" can represent the number of comments made by the students on that day.

In contrary, if the number of comments for student D is 22, then the mode "2" cannot represent the number of comments made by the students on that day because of the big difference between the mode "2" with two other values in the data which are 22 and 0. Therefore, although Lina had presented the knowledge of the mode as a form of data representation (P) but her justification towards utilising the mode as a form of data representation only when the mode and the data values are close to one another was found to be inappropriate (I).

In Task 4, Lina showed the idea of the mode as the average when she mentioned “almost all students”. She also showed the idea of the mean when she elaborated on the mean calculation. Lina said that the average can be based on the mean or the mode depending on the data. Regarding the information on the extreme data, she said that it can be ignored. She said that if the extreme data is not ignored then it might affect the average.

Lina said that the mode is the best measure to represent the data in this task. However, Lina did not mention if the median can represent the data when the median is supposed to be the best measure to represent data involving a population such as the Malaysian primary school students. Here the median is definitely a more robust measure. Therefore, Lina did not present the knowledge of the median as the average when it is supposed to be (NP). However, Lina’s presented the knowledge of the mode as the average and the mean as the average which were found not problematic given to proper justification because an average can be based by any of the three measures of central tendency (A).

In Task 5, Lina utilised the idea of median to obtain the weight of the median child. Lina said that the data has to be first arranged in ascending or descending order. Lina mentioned that the median for nine children is 19 and the median for all ten children after including 43 kg is also 19. However, Lina carried out the procedures incorrectly to obtain these median values. She elaborated on the even median procedure for the nine children which is supposed to be odd and the odd median procedure for all ten children which is supposed to be even. Therefore, Lina’s knowledge of the idea of the median was present (P) but incorrect (I).

Lina mentioned that the mean is not adequate to represent the weight of the ten children. She explained that the inadequacy of the mean is because of the tenth child’s weight which she said is an outlier. Lina said that the median would be a more suitable measure to represent the weight of all ten children. She added that, if the mean was considered to represent the weight of all ten children, then the weight of the tenth child has to be discarded first and then calculate the mean based on the remaining nine data. Otherwise, she said that the median would be better. Therefore, Lina's knowledge related to the robustness of the measures of central tendency was found to be present (P) and accurate (A).

In Task 6 and Task 7, Lina could construct data sets based on the given average value and the conditions imposed. In Task 6, Lina had utilised average as a typical value by placing the values for all seven bags of crisps as the average value. Lina justified this based on the "backward" mean calculation where she calculated the total of price for all seven bags from the given average and then divided with seven to obtain 27 for each bag. She also mentioned that placing 27 for all the seven bags of crisps is easier. Therefore, Lina's knowledge of constructing a data set based on the idea of average was present (P) and found to be appropriate (A).

In Task 7, Lina based her construction on random yet reasonable values that were around the average 27. She did not place numbers that were too far away from 27 because she knew that these can result in outliers. Lina again fell back on the backward mean calculation where she calculated the total price for all seven bags from the given average. This total was used to guide her construction where she made sure that the values that she placed for all seven bags of crisps were within this total. Lina's knowledge of constructing a data set based on the idea of average for Task 7 was found to be present (P) and appropriate (A).

Table 4.35

Coding rubrics for determining level of Lina's subject matter knowledge of measures of central tendency with reference to context

	Subject Matter Knowledge of Measures of Central Tendency with reference to Context																															
Task	Mean as average				Mode as average				Median as average				Idea of mode				Idea of median				Mode as data representation				Robustness of measures				Idea of average			
	P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP	
	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I
Task 1	X																															
Task 2								X					X								X											
Task 3													X									X										
Task 4											X																					
Task 5																		X							X							
Task 6																													X			
Task 7																													X			
	1										**1**		**2**								**1**				**1**				**2**			

Legend: P = Present NP = Not Present A = Appropriate I = Inappropriate

Lina's Percentage of Subject Matter Knowledge of Measures of Central Tendency with Reference to Context

$$= \frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)} \text{ X } 100\% = \frac{\mathbf{8}}{\mathbf{11}} \times \mathbf{100\%} = 72.73\% \text{ (Medium)}$$

Level of subject matter knowledge of measures of central tendency in handling bias. In Task 8, Lina calculated the mean for all nine data. She added up all the nine data and then divided the total with nine. She got the mean as 7.18. Therefore, Lina not only identified that the mean can be obtained but also summarised the data using the mean by calculating the mean value. Therefore, Lina's knowledge of identifying and summarising the data using the mean was found to be present (P) and correct (A).

Lina said that the median can be found after the data is arranged in ascending order. She said the median was 6.2 which was the middle number in the arrangement. Therefore, Lina's knowledge of identifying and summarising the data using the median was found to be present (P) and correct (A).

Lina also identified the mode as 6.3. She identified the mode based on the highest frequency where 6.3 had the highest frequency. Therefore, Lina's knowledge of identifying and summarising the data using the mode was found to be present (P) and correct (A).

Lina mentioned that in this task either the mode or the median are best suited measures to represent the data because these measures are around the majority of the data except for one particular value which is 15.3. Moreover, Lina mentioned that in this task, the mean is not so suitable because of the outlier 15.3 which was included in the mean calculation. Therefore, Lina's knowledge related to the best representation measure is found to be present (P) and appropriate (A).

Similarly, in Task 9, Lina calculated the mean for all five data. She added up all the five data and then divided the total with five. She got the mean as 76 800. Therefore, Lina not only identified that the mean can be obtained but also summarised the data using the mean by calculating the mean value. Therefore, Lina's knowledge of identifying and summarising the data using the mean was found to be present (P) and correct (A).

Lina said that the median can be found after the data was arranged in ascending order. She said the median was 54 000 which was the middle number in the arrangement. Therefore, Lina's knowledge of identifying and summarising the data using the median was found to be present (P) and correct (A).

Lina mentioned that in this task there is no mode. She explained that the frequency for each salary is one and there is no highest frequency here. So there is no mode in this task. Therefore, Lina's knowledge of identifying and summarising the data using the mode was found to be present (P) and correct (A).

Lina mentioned that the median is best measure to represent the data because the data values after ignoring 180 000 are around the median. Therefore, Lina's knowledge related to the best representation measure is found to be present (P) and appropriate (A).

Table 4.36*

Coding rubrics for determining level of Lina's subject matter knowledge of measures of central tendency in handling bias

Task	Subject Matter Knowledge of Measures of Central Tendency in Handling Bias															
	Identifies and summarises data using mean				Identifies and summarises data using median				Identifies and summarises data using mode				Best data representation measure			
	P		NP		P		NP		P		NP		P		NP	
	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I
8	X				X				X				X			
9	X				X				X				X			
	2				**2**				**2**				**2**			

Legend: P = Present NP = Not Present A = Appropriate I = Inappropriate

Lina's Percentage of Subject Matter Knowledge of Measures of Central Tendency in Handling Bias

$$= \frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)} \text{ X } 100\% = \frac{\mathbf{8}}{\mathbf{8}} \times \mathbf{100\%} = 100\% \text{ (High)}$$

Level of subject matter knowledge of measures of central tendency in problem solving. In Task 10, Lina carried out the backward mean calculation to obtain the total grade point average for College A based on two semesters as 6.4. She also found the total grade point average for College B based on three semesters as 11.4. Therefore, Lina's knowledge of performing the backward mean calculation to obtain the total from the mean is found to present (P) and correct (A).

Lina explained that the mean which was given for each college was for the number of semesters involved in that particular college. She also added the total grade points of both colleges and the total semesters of both colleges before she proceeded in the weighted mean calculation. Thus, Lina's knowledge related to the representative nature of the mean is found to be present (P) and appropriate (A).

Finally, Lina carried out the forward mean calculation to find the weighted mean. She divided the total grade points of both colleges with the total semesters of both colleges and found the weighted mean as 3.56. Therefore, Lina's knowledge of the forward mean calculation is found to present (P) and appropriate (A) despite the task not explicitly stating to find for the weighted mean.

In Task 11, Lina performed the backward mean calculation and obtained the total hours of TV watched per weekend for both rural and urban groups. She said that she multiplied the number of students in each group with the group average. Lina found the total hours of TV watched for rural group by multiplying 25 with 8. Similarly, for the urban group she multiplied 75 with 4. Thus, Lina's knowledge of performing backward mean calculation was found to be present (P) and appropriate (A).

Lina explained that she multiplied the average of each group with the number of students in each group because the given average was for the group. She also mentioned that the totals of both the groups were added and the total number of students of both groups were found before she proceeded with further calculation because she wanted to find the overall average for all the students. Thus, Lina's knowledge related to the representative nature of the mean is found to be present (P) and appropriate (A).

Finally, Lina carried out the forward mean calculation to find the weighted mean. She divided the total number of hours of both groups with the total number of students and found the average as 5. Therefore, Lina's knowledge of the forward mean calculation is found to present (P) and appropriate (A) despite the task not explicitly stating to find for the weighted mean.

Table 4.36

Coding rubrics for determining level of Lina's subject matter knowledge of measures of central tendency in problem solving

Task	Subject Matter Knowledge of Measures of Central Tendency in Problem Solving											
	Backward mean calculation				Representative nature of the mean				Forward mean calculation			
	P		NP		P		NP		P		NP	
	A	I	A	I	A	I	A	I	A	I	A	I
10	X				X				X			
11	X				X				X			
	2				**2**				**2**			

Legend: P = Present NP = Not Present A = Appropriate I = Inappropriate

Lina's Percentage of Subject Matter Knowledge of Measures of Central Tendency in Problem Solving

$$= \frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)} \text{ X } 100\% = \frac{6}{6} \times \mathbf{100}\% = 100\% \text{ (High)}$$

Level of subject matter knowledge of measures of central tendency in making inference. In Task 12, Lina calculated the mean scores for both the players by adding all the scores and divided the total with the number of games that they played. She found the median by arranging the data in ascending order and then estimated the median as the average of the two middle numbers in the arrangement. She also found the mode by looking at the most frequent occurred scored score.

Lina mentioned that Ramu's mean was around 19 to 20, his median was around 20 to 21 and his mode was 21. Whereas, she said that Beng's mean and median was around 21to 22 and his mode was 18. Therefore, Lina's knowledge of summarising equal sized numerical data using the measures of central tendency without the task explicitly stating to do so was found to be present (P) and appropriate (A).

Lina utilised all three measures of central tendency and looked at the consistency of these measures to decide on the player to be selected for the team. However, when Lina was probed to choose the player based on one measure, she chose to use the mode. Ramu was chosen because he had higher mode as compared to Beng. Lina mentioned that she would only use the mean when there is no mode when the mean was the most appropriate measure to be used and should have been her first choice.

The mean is the most suitable measure to compare and draw conclusions on two equal sized numerical data that deals with the performance of the players. This is because the mean takes in account every score and represents all the scores in the data set. Therefore, the mean describes better the performance of the players and drawing conclusions based on the mean score is the best. Thus, Lina's knowledge of utilising the appropriate measure of central tendency to compare and draw conclusion on two equal sized numerical data was found to be not present (NP) and also problematic (I)

In Task 13, Lina summarised the data given using all three measures of central tendency. Lina found that Ali's mean was 22.25, his mode was 25 and his median was 23.5. Whereas for Khan, she found that his mean was 23, his mode was 16 and his median was 25. Therefore, Lina's knowledge of summarising unequal sized numerical data using the measures of central tendency without the task explicitly stating to do so was found to be present (P) and correct (A).

Lina looked at the individual scores and the consistency of the scores to choose the player to be selected for the team. She said that she did not utilise any of the measures of central tendency to compare and draw conclusion based on the scores. However, when she was probed to use one measure, she mentioned that she would use the mode. Lina said that she would only consider the mean when there is no mode.

The mean is a meaningful real-world concept and a better measure to compare and conclude on unequal sized data. The mean has an advantage of representing all the data in the data set. Therefore, the mean is a better measure to compare and draw conclusion about the performance of the players. Thus, Lina's knowledge of utilising the appropriate measure of central tendency to compare and draw conclusion on two unequal sized numerical data was found to be not present (NP) and also problematic (I)

In Task 14, Lina utilised the measures of central tendency to summarise two equal sized graphical data without the task explicitly stating to do so. She found the mean, mode and median for both the classes. Therefore, Lina's knowledge of summarising two equal sized graphical data using the measures of central tendency was found to be present (P) and correct (A).

Lina utilised all three measures of central tendency to compare and decide on the class that performed better. She compared the mean, the mode and the median and found that 5A performed better than 5B because 5A had better mean, mode and median. However, when she was probed to use one measure, Lina mentioned that she would use the mean because usually the mean is used but later she changed her mind to the mode. According to her, the mode showed the highest frequency for students with corrects and would be a better measure to compare the performance of the classes.

Thus, Lina's knowledge of utilising the appropriate measure to compare and draw conclusion on two equal sized graphical data was found to be not present (NP). This is because the mean should have been Lina's choice to compare and draw conclusion about two equal sized data presented in graphical form. Lina's lack of knowledge in relation to the mean as a tool to compare and draw conclusion on two equal sized graphical data was found to be problematic too (I).

In Task 15, Lina utilised the mode to summarise two unequal sized graphical data. Therefore, Lina's knowledge of summarising two unequal sized graphical data using the measures of central tendency although she only used the mode was still marked to be present (P) and correct (A).

Lina compared the mode and the class that had the biggest mode was chosen as the class that performed better. Lina said that she did not prefer to use the median or the mean because the data looked skewed which might affect the mean and the median. Lina lacked in the knowledge that the mean is the most suitable measure to compare two unequal sized data that dealt with the performance. Therefore, Lina's knowledge of utilising the appropriate measure of central tendency to compare and draw conclusion about two unequal sized graphical data was found to be not present (NP) and also problematic (I).

Table 4.37

Coding rubrics for determining level of Lina's subject matter knowledge of measures of central tendency in making inference

	Subject Matter Knowledge of Measures of Central Tendency in Making Inference																															
Task	Summarises two equal sized numerical data using measures of central tendency				Utilises the appropriate measure to compare and draw conclusion about two equal sized numerical data				Summarises two unequal sized numerical data using measures of central tendency				Utilises the appropriate measure to compare and draw conclusion about two unequal sized numerical data				Summarises two equal sized graphical data using measures of central tendency				Utilises the appropriate measure to compare and draw conclusion about two equal sized graphical data				Summarises two unequal sized graphical data using measures of central tendency				Utilises the appropriate measure to compare and draw conclusion about two unequal sized graphical data			
	P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP	
	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I
Task 12	X							X																								
Task 13									X							X																
Task 14																	X							X								
Task 15																									X							X
	1								**1**								**1**								**1**							

Legend: P = Present NP = Not Present A = Appropriate I = Inappropriate

Lina's Percentage of Subject Matter Knowledge of Measures of Central Tendency in Making Inference

$= \frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)}$ X 100% = $\frac{\mathbf{4}}{\mathbf{8}} \times \mathbf{100\%}$ = 50% (Medium)

Lina's mean percentage of subject matter knowledge for all four constructs $\frac{\mathbf{72.73+100+100+50}}{\mathbf{4}} = \mathbf{80.68\%}$. Therefore, Lina's level of subject matter knowledge of measures of central tendency is high

Case Study Two: Bella

Bella was 23 years 4 months old when she was interviewed. At the point of data collection, she was pursuing a 4-year Bachelor of Science with Education (B.Sc. Ed.) program at a public university. She majored and minored in mathematics and chemistry respectively. Bella obtained A+ in Mathematics and Additional Mathematics in her SPM examination (equivalent to O level examination). She scored a B+ in Mathematics in the Matriculation examination.

At the time of data collection, Bella was in the final semester of fourth year studies. She attained 2.98 in the Cumulative Grade Point Average (CGPA) for her first three years of studies in the public university. She does not have any teaching experience prior to this interview. The following sections discuss the findings of Bella's subject matter knowledge of measures of central tendency in each of the four different constructs, namely with reference to context, in handling bias, in problem solving, and in making inference that were emerged from the clinical interview.

Types of subject matter knowledge.

Subject matter knowledge of measures of central tendency with reference to context.

Task 1. Bella's knowledge of the mean is average. This was clear in her explanation on how the average was obtained in which Bella elaborated the mean calculation. Bella knew that for the context of the data that was referring to the number of children, a whole number for the average is more appropriate. However, she said that the use of the mean algorithm can produce a decimal average.

Table 4.38

Excerpt B1

R: How do you think the average was obtained?
S: Take some samples because it is impossible for us to take data from all the Malaysian and look at how many children they have. It is difficult. So just take some samples maybe in thousands and then take average. This family got how many children and this family got how many children.
R: Can you explain?
S: Add all and divide.
R: Divide with what?
S: Divide with the total number of families taken.
R: What do you add up?
S: Add up all the number of children from all families and then divide by the number of families taken.
R: Can you show and explain using an example?
S: Let say family 1 has x_1, family 2 has x_2, until family n with x_n children and then we add up x_1, x_2, until x_n and then divide by n which is the total number of families.
R: They say that the families have 2.3 children. How can the average be 2.3, and not a counting number like 1, 2, 3, or 4?
S: Yes. Because it is referring to the number of children so if referring to humans is must be in whole number. But because we take sum of all the children and divide by number of families. A number divide by a number is not necessary to get a whole number. It can be a decimal. So, the average 2.3 is possible but maybe we can round off to 2 children to make it more realistic.

In Task 1, when Bella was asked on how the average would have been obtained, she started by explaining that the data is taken from a sample because it is difficult to get the data for the whole Malaysian population. Bella mentioned "add all and divide" in which she explained that to add the number of children for all the families and then divide with the number of families considered.

Bella further elaborated this in her example that if family 1 has x_1 children, family 2 has x_2 children until family n which has x_n children, to add all x_1, x_2 till x_n and then divide by n, the total number of families.

Bella explained that since the context of the data referred to the number of children therefore ideally the average has to be a whole number. However, she elaborated that a division of any two numbers will not necessarily result in a whole number. It can be a decimal number. Thus, according to Bella, the average 2.3 is possible but to make this value more realistic maybe it can be rounded off to 2. The above Excerpt B1 describes this.

Task 2. Bella utilised the concept of the highest number of demands to place the order for the shoe sizes in her shop. She also mentioned that the mode can be used a form of data representative but cannot be used as an average. In fact, Bella believed that the average is synonymous to the mean.

Table 4.39

Excerpt B2

R: How would you place the order for the shoe sizes in your shop?
S: I order the most for size 5 and then size 7 followed by 4 and 6 and lastly 8 and 9.
R: So how do you place the order?
S: Based on the highest number of demands.
R: If you want to choose one female shoe size, which shoe size would you choose?
S: Size 5
R: Why?
S: Because it has the highest demand among all the shoe sizes.
R: Can this chosen shoe size represent the female shoe size in your shop?
S: I think yes because according to data, there are a lot of people who wear size 5.
R: Why?
S: I think if I am not mistaken one of the way to represent we can use the mode. But it is better to take the average actually. But taking the mode is also one of the way. So to me it can be used to represent the female shoe size. Size 5 can represent the female shoe size.
R: What would you use as the average shoe size?
S: 5.85
R: How did you get this value?
S Since there are 5 who wears size 4, 12 size 5, and so on. Multiply 5 with 4, 12 with 5… and then add all the products. Then divide by the total number of data 33 so we get 5.85. Since maybe for the same reason, shoe size cannot be in decimal so I take the nearest one, round off to 6.
R: Would you consider 5 as the average shoe size?
S: No, I think I would not.
R: Why?
S: Because based on the average that I counted here I get 5.85 which is nearest to 6.
.
.
.
R: Can we use the mode as the average?
S: I think no. If we see the exact term, mode cannot be the mean. It is not average. But to represent we can use the mode. Since most use size 5 so I will order more size 5. Maybe if not in the context of my shop only but for all females than we have to take the data from other shops also. Then the average should be considered because it will be more accurate.

Bella mentioned that she would place the order for the shoe sizes in her shop based on the highest number of demands. She said that it would be the most for size 5 followed by sizes 7, 4, 6, 8, and 9. Bella said that she would choose size 5 because it has the highest demand among all the shoe sizes. In fact, she mentioned that size 5 can represent the shoe sizes in her shop because “a lot of people” wore size 5.

When she was probed further, she mentioned that the mode can be used to represent the data in this situation. However, she also mentioned that the average is a better measure to represent data. Nevertheless, in the given situation, Bella felt that the mode size 5 can be used to represent the female shoe sizes in her shop because the mode can be one way to represent the data.

When Bella was asked on what she would use as the average shoe size, she said 5.85. She mentioned that 5 customers wore size 4, 12 size 5 and so. She elaborated that to multiply 5 with 4, 12 with 5 and so. Then add all the products obtained and divide with the total number of data which is 33 to get 5.85. Bella also added that since the context of the data deals with shoe sizes therefore the value cannot be not presented in decimal number. She said that she would round off the value 5.85 to the nearest which is 6.

Bella mentioned that she would not consider 5 as the average shoe size because the average that she counted was 5.85 in which the nearest whole number is 6. However, when Bella was asked if she would use the mode as the average, she mentioned "no". She explained that according to the exact definition of terms, the mode is not the mean or the average. However, Bella explained that the mode can be utilised to represent data. She said that since most of her customers used size 5 therefore she would order more on size 5. However, she also mentioned if it is not involving her shop but the data from other shops too, then the average should be considered because it is more accurate. Excerpt B2 illustrates Bella's points on this.

Task 3. Bella identified the mode as the highest number of comments. She mentioned that the mode is the highest frequency. Bella also mentioned that she would only utilise the mode to represent data if the calculated average value for the data is also around the same value as the mode.

Table 4.40

Excerpt B3

R: What is the mode?
S: Mode is 22
R: How did you obtain this value?
S: Because that is the highest number of comments among those students. So, it is 22. Because what I understand about the mode is the highest frequency so based on the table, student D gives 22 comments while the others gave at most 5. So, 22 is the mode.
R: If you were to represent one value for the number of comments made by the students on that day, would you use this value?
S: No
R: Why?
S: It is much different compared to the other comments and based on what I have learned if the data is much different than the others we must do a test to determine whether we can take that data or not. It is an outlier or not. If it is than we can discard it.
R: If I were to change the value 22 to 6, what will the mode be now?
S: 6
R: How did you get this value?
S: The highest number of comments
R: Would you use this value now to represent the data?
S: I must find the average first if it is 6 than yes can. I add the total up everything 5 plus 2 plus 6…. and divide by the number of students. If I get 6 than I say yes.
R: What if you do not get 6?
S: I take the average. And for the same reason like the previous, this is comments so it cannot be in decimal, so I must round off to whole number.

Bella identified the mode as 22. She mentioned that the mode is highest number of comments among those students. She elaborated that according to her understanding the mode is the highest frequency, so based on the table, student D gave 22 comments as compared to the others who gave at most 5. Thus, Bella claimed the mode is 22.

Bella mentioned that she would not use this value to represent the number of comments made by the students on that day. She explained that the mode is far from the other values given in the table. She elaborated that if the data point is different than the others, a test has to be carried out to check if the data is an outlier which can be discarded.

The researcher probed Bella further by changing the number of comments 22 to 6 in which she was asked to identify the mode again. Bella mentioned that the mode is now 6. Again, she based this on the highest number of comments. Bella mentioned that she would only use this mode to represent the number of comments made by students on that day based on the calculated average. She elaborated that if she obtains the average as 6 then she would use the mode to represent the data. Otherwise, she would just take the average itself to represent the data after rounding off the decimal value to a whole number because the number of comments cannot be in decimal. The above Excerpt B3 illustrates Bella's points on this.

Task 4. Bella mentioned that the average can mean the most. However, she elaborated on the mean calculation when she was asked to explain on how the average was obtained. Bella said that to calculate an average, all that data need to be included.

Table 4.41

Excerpt B4

R: What does average mean in this sentence?
S: Based on my understanding, the average is that most of the students watched 3 hours of TV per day.
R: How do you think they got this average?
S: Say that student 1 watched x_1 hours, student 2 watched x_2 hours, and so till student n watched x_n hours. Add all the number of hours and then divide with the number of students.
R: What about 12 hours of TV per day there?
S: [Paused awhile] if there is a group of students who watched more than 12 hours TV per day than there should be a group who watched TV maybe less than 3 hours. A group that does not watch much. That is why we get 3 hours of TV per day as the average.
R: Can you explain using an example?
S: Based on the formula to find the average. You must take all the hours watched by all the students and then divide by the total students.
.
.
.
R: So, will the information on a small group of students watched more than 12 hours of TV per day be included?
S: Yes.
R: How?
S: Just include in the calculation
R: Do you think it will influence the average?
S: Yes
R: How?
S: Because a group of students which means one or two. It does affect the average value.
R: Beside this, is there any other way you can obtain average?
S: I think based on the information, maybe.
R: How?

S: Maybe can consider the mode also. No no no... I think no. Can you repeat again?
R: Beside the one that you told me which is based on the calculation, is there any other way to obtain the average?
S: I think no. Just take the average.
R: Can you explain on the 12 hours?
S: Yes. Since here there is a group of students who watched more than 12 hours while the average is 3, which is I think is quite a big difference. So, what I understand is that there must be also a group of students that watched less than 3 hours. That is why we get the average as 3 hours. So, a group watched more than 12 hours and if there is a group watched 7, 8 so we cannot get average 3. So, the included numbers must be small enough to get the average as 3 hours.

Bella mentioned that according to her understanding, the average "is that most of the students watched 3 hours of TV per day". Bella elaborated on the mean calculation when she was asked on how the average would have been obtained. She mentioned that if student 1 watched x_1 hours, student 2 watched x_2 hours and till student n that watched x_n hours, all the number of hours is added and then divided with the number of students.

When Bella was asked to explain on the information that a group of students watched more than 12 hours of TV per day, she paused awhile. She explained that if there was a group of students who watched more than 12 hours TV per day than there should also be another group who watched maybe less than 3 hours which can result in the average as 3 hours of TV per day.

Bella explained that based on the formula to find the average, the total number of hours of TV watched by all the students is divided with the total students. Bella mentioned that the information on a group of students who watched more than 12 hours of TV will also be included in the calculation. She also said that this information will affect the average value.

Bella mentioned that there might be other ways to find average. She mentioned that the mode can be considered. However, she became doubtful and then changed her response to no other ways to find the average but the average itself. Bella was probed again on the information on 12 hours. After when she explained clearly that 12 hours with the average 3 hours has quite a big difference. Therefore, according to her understanding there will be another group that watched less than 3 hours. The values in this group are small enough to average out to 3 hours which might not be possible if there were more in 7 or 8 hours in the group. The above Excerpt B4 illustrates this.

Task 5. Bella found the median by rearranging all the weights from the smallest to the biggest. Then she added all the values and divided the total with two. The value that she obtained from the division was checked for its place position in the ordered arrangement by adding the arranged weights one by one till she arrived at the position where this value might be positioned. Using this position, she obtained the median by averaging the two weights that fell in between the position. Bella also mentioned that the mean is not adequate to represent the weight of all ten children. Instead she suggested to find the mean of the nine children after 43 kg was excluded and this mean can represent the weight of all ten children.

Table 4.42

Excerpt B5

R: Which is the weight for the median child?
S: 19
R: How did you get this?
S: I think, I add all the weight in kilograms and divide by two. I get 89.5. Before that I rearrange back the weight, start from the lowest to the highest and then I add all the weights and divide by two. I get 89.5 and then I add up from the left, from the arranged sequence and find for the position for 89.5 which is at 19. I add this number (points all the numbers in the sequence, until here I get 85 and then I add 19 more, I found that 89.5 is somewhere in the middle here. So, to get the median, since this value (points to 89.5) is in between these two weights so we have to take the average of this two weights. So, we get 19.
R: Which is the median if we include another child who weighs 43 kg?
S: If add another child with 43 kg then I find the total first I get 222 and then divide by 2. I get 111. Then I find the position for 111 which I find it is here between 19 and 24. So to get the median, we find the average of these two weights which is 19 plus 24 divide by 2, we get 21.5.
R: Is there any other ways that you can find the median besides the one you have used here?
S: I think no. I do not know any other ways to find the median unless this one.
R: Is it adequate to use mean to represent the weight of the 10 children?
S: Mmm… I think no.
R: Why?
S: Because the tenth child is much heavier compared to the other children.
R: Then how do you deal with this?
S: Just take the weight of the earlier nine children.
R: What about the 43 kg?
S: Just ignore because it is far away from the rest nearly twice more than the rest. Find the mean for nine children.

R: Besides this, is there any other way to represent the weight of the 10 children?
S: I think no. I will use mean.
R: How do you find the mean?
S: Add all the weight for these nine children and then divide by nine. Ignore the 43 kg because it is far away from the other data.

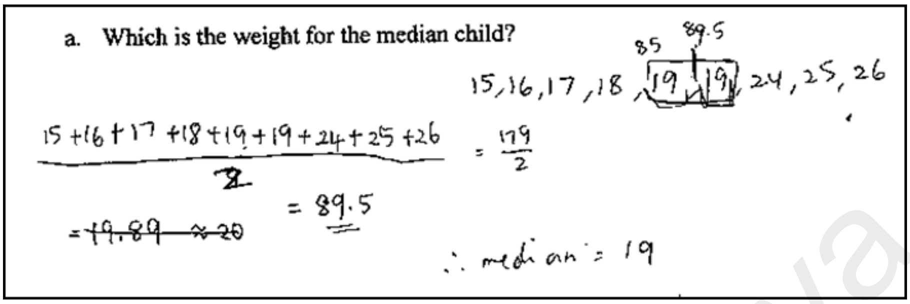


Figure 4.8. Entry B1

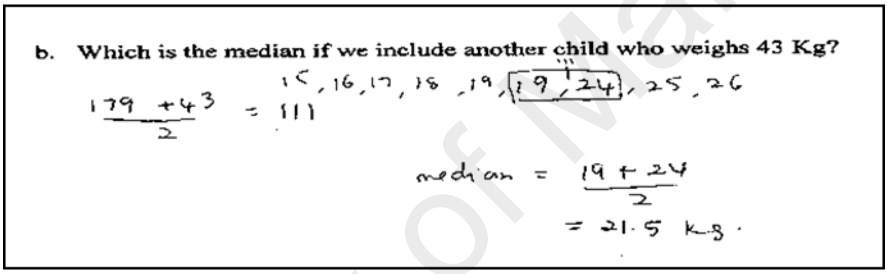


Figure 4.9. Entry B2

Bella identified the median by first arranging all the weights from the smallest to the biggest. Then she added all the weights for the nine children and divided the total with two in which she got 89.5. Bella added one by one the weights starting from the smallest weight in the rearrangement. She added the weights till she arrived with the total somewhere near to 89.5. She pointed at the first 19 and mentioned that the total up till this number is 85. She elaborated that if she added the subsequent 19, then the total will be more than 89.5. So according to Bella, the total 89.5 which gives the median is in between the first and second 19. Therefore, she said that to average these two weights and the average of 19 and 19 is also 19 which is the median. This was clearly described in the above Excerpt B5 and also Entry B1.

Bella elaborated on a similar procedure to obtain the median for all ten children after the weight of another child which is 43 kg was included. She arranged all the ten weights from the smallest to the biggest. Then she added all the ten weights in which she got as 222. This total was later divided by two which gave her 111. She added the weights in the rearrangement from the smallest one by one until she arrived at the total 111 or somewhere near to 111. She found that the total 111 will be positioned in between the weights 19 and 24. Therefore, Bella added these two numbers and divided the total with two. She got the median as 21.5. The above Excerpt B5 and Entry B2 illustrates Bella's points on this.

Bella was not aware of any other ways to find the median except the one that she elaborated above. According to Bella also, the mean is not adequate to represent the weight of all ten children. She explained that this is because of the weight of the tenth child which is much heavier than the weights of the other children. Bella mentioned to just consider the weight of the nine children and ignore 43 kg because this weight is nearly twice heavier than the rest of the weights. She said that she would just find the mean for the nine children after 43 kg is discarded. Bella also said that she is not aware of other ways to represent the weight of all ten children except by using the mean. The above Excerpt B5 describes this.

Task 6 and task 7. Bella placed the price on each bag of crisp randomly but around the average price. She calculated the total price for all the bags by using the backward mean calculation and utilised this total as a guide in the construction.

Table 4.43

Excerpt B6

R: How did you place the price on each of the bags?
S: Just put.
R: How?
S: Since the average is 27 for seven bags so first I times 27 by 7 to get the total price for all the seven bags. I get 189. So, I randomly place one by one.
R: How did you randomly choose these values?

S: Mmmmm….at first, I choose one value or price, I minus this value from the total. Then choose the second value and minus again from the balance total and so on…. Until the total that has been subtracted is 189 to give a balance of zero.
R: How did you base the values?
S: I based upon the average value. Because [looked unsure] actually there are possibilities that some of the values are far away from the average but to make it convenient for me I chose just the values that are close to the average. There are many possibilities. This is one of the many possibilities where the prices of the seven bags are very close to the average. There are also possibilities where some of the bags say three that might have values that are much higher than the average price.

Bella mentioned that she “just put” the prices for the bags of crisps. She elaborated that since the average is 27 for all seven bags, therefore 27 multiplied with 7 gave her the total price for all seven bags as 189. She mentioned that she randomly placed the prices one by one. When she was probed on how she randomly chose the values, she explained that she first chose one value, she subtracted this value from the total, and then she chose the second value and minus again from the total and so on. She mentioned that she will do this until she gets all the seven values and she gets the balance from the total 189 as zero.

Bella mentioned that the values were based on the given average. She elaborated that she can come up with many possible values, some can be far from the average but for her convenience she chose values that are close to the average. Her construction here is one of the many possibilities where the prices are very close to the average. She added that another possible construction can be by having three of the prices much higher than the average. The above Excerpt B6 illustrates Bella’s points on this.

Table 4.44

Excerpt B7

R: Look at task 7. Task 7 says that you cannot place 27 as the price of a bag of crisps. How will you deal with this task?
S: Just the same as the earlier task. I take randomly the first value or price and then minus with the total 189. Then the second value minus from the balance total and so on…repeat until the whole total is subtracted off to zero.
R: Any other ways that you can base the values?

S: Actually, this is also very close to the average value. You can take different values. You can take some bigger values than 27. But I think taking bigger values will make my calculation difficult. That is why I take values which are closer to 27. R: How did you base these values? S: I based on the total for the seven bags and each one value I placed a random value.

Similar to the earlier task, Bella placed random values as the price values for the seven bags. She elaborated that she took the first random value and subtracted this value from the total price 189. She repeated the same with the second value in which she subtracted this value from the balance total. She repeated this until the whole total is subtracted off to zero. Bella mentioned that the values that she had taken are close to the average. She said that different values can be considered. She added that some bigger values than 27 can be considered but she felt this will make her calculation difficult. Therefore, this was the reason why she preferred to take values which are closer to 27. Bella mentioned that she had based the random values for the prices on the total price for all seven bags of crisps. Excerpt B7 describes Bella's points on this.

Summary.

Table 4.45

Subject matter knowledge with reference to context

Task	Statistical element	Conclusion
Task 1	Mean as average	Bella's knowledge of the mean is average. Bella mentioned that the average is obtained by using "all add and divide". Bella elaborated on the mean calculation in her example too. Bella mentioned that for the context of the data that was referring to the number of children, a whole number for the average is more appropriate. However, she said that the division of any two numbers in the mean calculation can produce a decimal average. She said that the decimal average can be made realistic given to the context if it was rounded off.
Task 2	Idea of mode	Bella utilised the idea of mode without the task explicitly stating to do so. She utilised the frequency tally to find the frequency of each shoe size and then she identified the mode based on this as the shoe size with the highest number of demands.
	Mode as data representation	Bella said that the mode can represent the shoe sizes in her shop because of its highest demand. She said that one way to represent any data is by using the mode. However, she claimed that using the average is better.

	Mode as average	Bella did not utilise the mode as the average. Instead she calculated the average using the mean algorithm. She got a decimal average in which she said should be rounded off to a whole number given to the context of the data. Bella said that according to the definition, the mode is different from the mean or the average. Therefore, she mentioned that the average cannot be based on the mode.
Task 3	Idea of mode	Bella identified the mode as the highest number of comments in which she said that the mode is the highest frequency
	Mode as data representation	Bella mentioned that she would not use the mode to represent the data. She claimed that the mode is different than the rest of the data. However, Bella said that she would use the mode value to represent the data if the average that she calculated is the same as the mode. Otherwise she said that she would take the average to represent the data.
Task 4	Median as average	Bella mentioned that the average means the most. However, her later explanations revealed that her knowledge of the average is the mean in which she elaborated on the mean calculation in her example. Bella said that in the calculation of the average, all the data including extreme data need to be considered.
Task 5	Idea of median	Bella found the median by rearranging all the weights from the smallest to the biggest. Then she added all the values and divided the total with two. The value that she obtained from the division was checked for its place position in the ordered arrangement by adding the arranged weights one by one till she arrived at the position where this value might be positioned. Using this position, she obtained the median by averaging the two weights that fell in between the position.
	Robustness of measures	Bella said that the mean is not adequate to represent the data that contained an outlier. Instead she said that the mean which is calculated after excluding the outlier can be considered. She did not mention of any other measures that can be used to represent the data.
Task 6	Idea of average	Bella calculated the total price using the backward mean calculation. Although she placed the prices randomly but she said she utilised the total in which the first random value was subtracted from the total followed by the rest one by one until all seven values and the balance from the total becomes zero. Bella mentioned that she based these random values quite close to the average.
Task 7	Idea of average	Similar to the earlier task, Bella calculated the total price using the backward mean calculation. Although she placed the prices randomly but she said she utilised the total in which the first random value was subtracted from the total followed by the rest one by one until all seven values and the balance from the total becomes zero. Bella mentioned that she based these random values quite close to the average.

Subject matter knowledge of measures of central tendency in handling bias.

Task 8 and task 9. Bella identified the mode based on the most frequent occurring value and also the biggest value. She also identified the median using a value that she obtained by adding all the data and dividing the total with two. Then the position of this value is located using the cumulative total starting from the lowest value in the data. However, Bella was not consistent with this method when she dealt with Task 9. She neither arranged the data according to any particular order nor did she locate the position of this value using the cumulative total starting from the lowest value in the data.

Table 4.46

Excerpt B8

R: What are the measures of central tendency that can be obtained based on the above data?
S: Mode which is the highest frequency 15.3, mean is 7.18 and median is 6.25.
R: How did you get the median?
S: Arrange from the lowest value to the highest value. And then add all up and divide by two. I get 32.325. I find the position in the cumulative value. 32.325 is between 6.2 and 6.3. So, the median is the average of this two numbers which is 6.25.
R: How about the mean?
S: Mean, I total up everything and then divide by nine since there is nine students. So, I get 7.18.
R: How about the mode?
S: Mode is the highest frequency which is15.3. I think my mode is wrong. Mode is based on number of students so must be 6.3 because three students got 6.3.
R: Which is the best measure of central tendency to represent the actual weight of this object?
S: I think mean
R: Why?
S: Because it considers all the data.
R: How?
S: The data is showing the weight. The values are around 6 to 7 and only one value which is far 15.3. The mean considers all the data. Mean uses all the data so it is the best measure to represent the actual weight.

Bella mentioned that for the data in Task 8, the mode which is the highest frequency is 15.3, the mean is 7.18 and the median is 6.25. She explained that the median was obtained by first arranging all the data from the lowest to the highest value. Then she mentioned that she added all the data and then divided the total with two. She said that she got 32.325 in which she used this value to determine the position of the median. The position was determined by using the cumulative total starting from the lowest value. She found that 32.325 is positioned between the data points 6.2 and 6.3. Thus, she obtained the median as 6.25 by averaging these two numbers.

Bella explained that in order to get the mean, she added all the nine data and then divided the total with nine and got the mean as 7.18. Whereas the mode is 15.3, she said is the highest frequency. However, Bella realized that the mode that she obtained was incorrect in which she quickly corrected it to 6.3 which she claimed now was based on three students who got 6.3.

According to Bella, the mean is the best measure to represent the actual weight of the object. She said that the mean considers all the data. She explained that the data is on the weight of an object and most of the data is around 6 to 7 except for one far value which is 15.3. She claimed that the mean uses all the data, therefore it is the best measure to represent the actual weight. The above Excerpt B8 illustrates Bella's points on this.

Table 4.47

Excerpt B9

R: What are the measures of central tendency that you can obtain based on the above data?
S: Mode 180 000 because it is the highest salary. Mean is 76 800 which I total up all the salaries and then divide by five to get this value. But the median I am not really sure but I got 48 000.
R: How did you obtain the median as 48 000?
S: First I totalled up all the salaries and then I divided the total with two. I got 192 000. I found the cumulative salary and tried to locate the position of the cumulative salaries till 192 000 which is within the salary 48 000. So, the median is 48 000.
R: How about the mode?
S: The mode is the highest salary which is 180 000
R: Which is the best measure of central tendency to represent the actual salary of the company?
S: I think mean.
R: Why?
S: Same reason because consider all data. For the big value, do the test and determine whether the value can be accepted or rejected. If accepted the mean value is the one that I mentioned. If rejected than we take the average for four salaries.

For Task 9 Bella identified the mode as 180 000 which she based on the highest salary. She got the mean as 76 800 in which she said that she added all the salaries and then divided the total with five. However, when it came to the median, Bella claimed that she was not so sure. She said that she got the median as 48 000 and she elaborated that she first totalled up all the salaries. After when she divided the total with two and she got 192 000. 192 000 was used to locate the position of the median by using the cumulative salaries. She found that 192 000 is positioned within the salary of 48 000. Therefore, she concluded that the median is 48 000.

According to Bella, the mean is the best measure to represent the actual salary of the company. She explained that the mean considers all the data. Nevertheless, here Bella mentioned that for the big value which she referred to 180 000, a test need to be carried out to determine if this value can be accepted or rejected. If accepted than the mean will be the one that she mentioned earlier. On the other hand, if rejected than she said that the average should be based on the remaining four salaries. This is shown in the above Excerpt B9.

Summary.

Table 4.48

Subject matter knowledge in handling bias

Task	Statistical element	Conclusion
Task 8	Identifies and summarises data using mean	Bella found the mean by adding all the nine data and then divided the total with nine.
	Identifies and summarises data using median	Bella identified the median using a value that she obtained by adding all the data and dividing the total with two. Then the position of this value is located using the cumulative total starting from the lowest value in the data arranged in ascending order.
	Identifies and summarises data using mode	Bella initially identified the mode based on the biggest value in the data in which she believed was the highest frequency. However, later she based the mode on the most frequent occurring value.
	Best data representation measure	Bella mentioned that the mean is the best measure to represent the data because it considers all the data. She said that most data are around 6 to 7 and only one data was far from the rest, therefore the mean considers all and would be the best representation measure.
Task 9	Identifies and summarises data using mean	Bella found the mean by adding all the five data and then divided the total with five.

Identifies and summarises data using median	Bella was not very consistent with the procedure to find the median. She got a value from the total of all the salaries which was divided with two. However, she neither arranged the data according to any particular order nor did she locate the position of the value using the cumulative total starting from the lowest value in the data like what she did in Task 8.
Identifies and summarises data using mode	Bella based the mode on the biggest value instead of the most frequent occurring value here which was different than what she did with Task 8.
Best data representation measure	Bella mentioned that the mean is the best measure to represent the data because it considers all the data. However, here she mentioned that the extreme data needed to be tested first to see if it is to be accepted or rejected from the calculation of the mean. According to her, if the extreme data is to be rejected than the mean is based on the remaining data.

Subject matter knowledge of measures of central tendency in problem solving.

Task 10. Bella initial was confused on how to get the overall GPA. She got the overall GPA by adding both the GPAs for college A and college B and then she divided the total with two. However, after thinking and looking back at the problem, she found the overall GPA using a different approach in which she considered the number of semesters involved in both the colleges into her calculation.

Table 4.49

Excerpt B10

R: What was the student's overall GPA?
S: I am not very sure but I have done something. I got 3.5. I just take the average for the two GPA.
R: Can you explain?
S: Because the first time, the student got 3.2 and the second time got 3.8. So, take average to get the overall GPA for this student.
R: How about the semesters involved?
S: I have no idea about the two and three semesters. I do not know how to use. [Paused for a very long time and was thinking carefully about the question and was trying to solve the question in the task sheet]. Then it is 3.56.
R: How did you get 3.56?
S: When the student enters College A for two semesters, he gets 3.2 GPA. So, for GPA for first semester and second semester are added and divided by two to get 3.2. When the student enters College B, he gets 3.8 for three semesters. So, the GPA for first semester, second and third semester we add up and divide by three we get 3.8. So, to get the overall, we need to find the total for all five semesters and divide by five. So, 3.2 times 2 and 3.8 times 3 is add up and divide by five for five semesters.
R: Why divide by five?
S: The total semesters is 5.

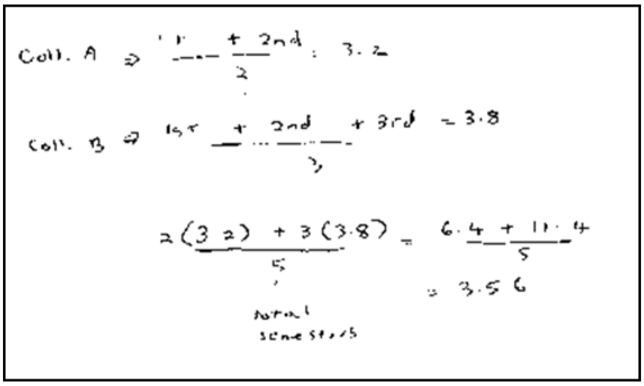


Figure 4.10. Entry B3

Initially Bella sounded not very sure on how to find the overall GPA. However, she revealed that she had done something and got the overall GPA as 3.5. She explained that she averaged the two given GPAs. However, after she was probed on the semesters, Bella started to think. She mentioned of not having any idea about the semesters and not knowing how to use the information. She paused for a very long time and gave it some thought while trying to work on the solution. After when Bella disclosed that the overall GPA is 3.56. This is also shown in Entry B3.

Bella explained that in college A the student spent two semesters and got a GPA of 3.2. So, she said that for the first semester his grade point is 3.2 and same goes for the second semester another 3.2. These two grade points are added and divided by two to give the GPA as 3.2. Similarly, she said that when the student entered college B, his GPA for three semesters are 3.8. Thus, his grade points for the first, second and third semesters added up and then divided with three to get the GPA as 3.8.

Bella elaborated that to get the overall GPA, the total grade points for all five semesters is divided with five. Therefore, she said that 3.2 should be multiplied with 2 and 3.8 should be multiplied with 3. Both the products are added up and then divided with five because a total of five semesters involved. The above Excerpt B10 describes Bella’s points on this.

Task 11. Bella found the total TV viewing time for both rural and urban groups by utilising the given group average. Later she calculated the average TV viewing time for all 100 students by adding both group totals to get the overall total and then divided this overall total with 100.

Table 4.50

Excerpt B11

R: Show clearly the average TV viewing time per weekend for all 100 students. Can you explain?
S: Since the average for 25 rural students is 8 and for 75 urban students is 4. To get the total hours for both rural and urban, 25 times 8 and 75 times 4 and then total up.
R: Why 25 times 8?
S: Because for 25 rural students, the average is 8 hours. I have to find the total hours watched by rural students. The same for urban students, 75 urban students watch an average of 4 hours. So, 75 times 4 gives me the total hours watched by urban students.
R: Why are you dividing by 100?
S: Divide by 100 to get the average for 100 students. 25 plus 75. I have 500 hours as the total hours watched by both rural and urban. I divided this total with 100 students. So, the average is 5 hours of TV per weekend for each student.

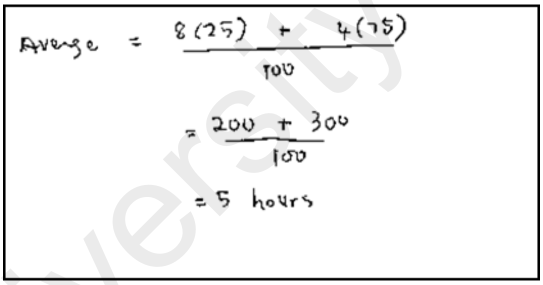


Figure 4.11. Entry B4

Bella mentioned that the average for 25 rural students is 8 and the average for 75 urban students is 4. Therefore, she said that to get the total hours for both rural and urban students, she multiplied 25 with 8 and 75 with 4. Then she totalled up both the products. Bella explained that the average 8 hours is for 25 students, so to get the total hours for all 25 rural students, she multiplied 25 with 8. She mentioned that the same for the urban students where the average 4 is for 75 urban students. Therefore, she multiplied 75 with 4 to get the total hours for urban students.

Bella mentioned that she divided the overall total hours watched by both rural and urban students with 100 because she wanted to get the average for all 100 students. She said that she got the total hours as 500 for both rural and urban students in which she divided with 100 students. Thus, she calculated the average as 5 hours of TV per weekend for each student. This is also shown in Entry B4. The above Excerpt B11 describes Bella's points on this.

Summary.

Table 4.51

Subject matter knowledge in problem solving

Task	Statistical element	Conclusion
Task 10	Backward mean calculation	Bella solved for the weighted mean for both the tasks without the task explicitly stating to do so. She had carried out the backward mean calculation, where she obtained the total for each group of data set from the given group means.
	Representative nature of the mean	Bella used the knowledge related to the representative nature of the mean in making sense of why data sets need to be combined in the weighted mean.
	Forward mean calculation	Bella also carried out the "forward" mean calculation to calculate the weighted mean.
Task 11	Backward mean calculation	Bella solved for the weighted mean for both the tasks without the task explicitly stating to do so. She had carried out the "backward" mean calculation, where she obtained the total for each group of data set from the given group means.
	Representative nature of the mean	Bella used the knowledge related to the representative nature of the mean in making sense of why data sets need to be combined in the weighted mean.
	Forward mean calculation	Bella also carried out the "forward" mean calculation to calculate the weighted mean.

Subject matter knowledge of measures of central tendency in making inference.

Task 12. Bella mentioned that she would recommend the player to be selected based on the average. She also mentioned that a one to one score comparisons on each set of game score can also be carried out and the player that won most of the games can be selected. However, Bella claimed that among these two methods, considering the average is better because she believed that the one to one score comparisons might not be accurate.

Table 4.52

Excerpt B12

R: If Coach Ahmad can only select one of the two players, who would you recommend he select?
S: I would recommend Beng.
R: Why?
S: Because based on the average that I get, he gets a higher average compared to Ramu. So, means that Beng has better performance.
R: How did you find the average?
S: The same procedure, for each of them, I total up all the scores that they get and then divide by six because there are six scores. For Ramu I get 19.67. For Beng 21.83. Since Beng has higher average so I will recommend him.
R: Is there any other ways that you can use to decide which player to choose?
S: Maybe I can compare each set of game score. First game set score Beng got better, second also, fifth and sixth. So Beng did better in four game sets as compared to Ramu which is only two. So, we can see that Beng did better than Ramu.
.
.
.
R: Which is the better method to select the player?
S: Average is better.
R: Why?
S: Because it considers all the scores. Sometimes one to one comparison might not be correct because maybe the player scored with little difference and maybe he lost with lot of difference. So, average is still the best.

Bella mentioned that she would recommend Beng to be selected for the team. She said that she based her decision on the average. She found that Beng got a higher average than Ramu, so she claimed that Beng had a better performance. Bella calculated the average by adding all the scores for each of the player and then divided the total with six because each of the player played six games. According to Bella, Ramu's average was 19.67 and Beng's average was 21.83. Since Beng's average was higher than Ramu's, therefore she decided to recommend Beng to be selected for the team.

Bella said another way that can also be used to select the player is by comparing each set of game scores. She found that Beng scored better in the first, second, fifth and sixth game sets as compared to Ramu who only scored the third and fourth game sets. Beng scored four game sets whereas Ramu only scored two game sets. Therefore, Bella saw Beng performed better than Ramu.

However, among the two methods that Bella mentioned, she considered the average as the better method to select the player. For Bella, the average considers all the scores of the player. She explained that the one to one game set comparison might not be accurate because maybe the player that won a particular set with marginal difference in the scores and he might have lost certain game sets with big difference in the scores. Therefore, she believed that this does not reflect well the player's actual performance and the average is still the best to be considered. The above Excerpt B12 describes Bella's points on this.

Task 13. Bella mentioned that for the situation in this task, only the average comparison can be used to determine the player to be selected. According to her, a one to one game set comparison cannot be done here because the players did not play equal number of game sets.

Table 4.53

Excerpt B13

R: Who would you recommend him to select?
S: Khan.
R: Why?
S: Because based on the average, Khan gets a higher average 23 while Ali 22.25.
R: How did you find the average?
S: I add up all the scores and then for Ali, I divide by eight because there are eight scores and for Khan, I divide by six because there six scores. Khan got a bit higher average than Ali. So, I would recommend Khan.
R: If you look at the data, the data is unequal sized. One player got eight scores and the other got six scores. Do you think you can use average?
S: Yes.

R: Why?
S: Because it is fair. All the games that the player played are taken count.
R: Is there any other way to determine the player to be selected?
S: I think here we cannot have compared the game sets because the number of games played is different for both Ali and Khan. Ali played 8 games while Khan played 6. In this case, like what I mentioned earlier, do comparison of average.

Bella mentioned that she would recommend Khan. She based her decision on the average where Khan got a higher average which was 23 as compared to Ali's average which was 22.25. She said that she found the average by adding all the scores for each of the player and then divided with the number of game sets that he played. She explained that for Ali, the total scores were divided with eight because he played eight games and for Khan, the total scores were divided with six because he played only six games.

Bella mentioned that the average can be used to compare and decide on the player to be selected because the average is a fair measure and takes count all the scores for the player. Bella added that in this situation a one to one game set comparison cannot be done because of unequal number of game sets involved for both the players. Therefore, she said that the comparison of the average scores is the most ideal way to select the player for the team. This is described in the above Excerpt B13.

Task 14. Bella had two different ways of identifying the player that she would recommend to be selected for the team. She said it can be based on the average score comparison or by looking at the graph. However, Bella preferred the average score comparison because she claimed it is clearer and accurate.

Table 4.54

Excerpt B14

R: Based on the graphs below, which class did better?
S: 5A did better.
R: Why?
S: Because the average score is higher.
R: How did you get the average score?
S: Since 1 student gets 3 corrects, 2 students get 4 corrects, 3 students get 5 corrects and 6 corrects, so times the number of students with the number of corrects in order to get the total scores. So, I

add up all the products 1 x 3 + 2 x 4 + 3 x 5 + 3 x 6 then I divide by the number of students. The total scores are 44 and the total student is 9. So, I divided the total scores with the total students to get the average as 4.89 for 5A. For 5B, the same way, 38 divide 9 and I get the average as 4.22. So 5A got higher average as compared to 5B. 5A did better.
R: Is there any other way to determine which class did better?
S: Mmm… by looking at the graph itself. For 5A, there are three students got six corrects, whereas 5B only got one student who got six corrects. And then for five corrects, 5A got three students whereas 5B got two students. Can see that 5A students scored more corrects than 5B.
R: You mentioned two ways to determine the player to be selected. Which one would you use?
S: I would use the average.
R: Why?
S: More accurate. The value shows clearly the class that did better.

Bella mentioned that 5A did better because the class average score is higher than 5B. She explained that the average score was obtained by first multiplying the number of students with the respective number of corrects. Bella mentioned that for 5A, 1 student got 3 corrects so 1 times 3, 2 students got 4 corrects so 2 times 4, 3 students got 5 corrects so 3 times 5 and so on. She added up all the products and then divided the total with the total number of students. According to Bella, 5A had a total score of 44 which was divided with 9 students to get the average as 4.89. On the other hand, she said 5B had a total score of 38 which was divided with 9 students to get the average as 4.22. Since the average score of 5A was higher than 5B, she said that 5A performed better.

Bella said that the performance of the class can also be determined by looking the graphs. Bella did a comparison on both the classes. She said that 5A had three students who got six corrects whereas 5B had only one who got six corrects. She compared again for five corrects where she found that 5A had three students whereas 5B had only two students. Therefore, Bella concluded that 5A students scored more corrects than 5B. However, she said that out of the two ways to determine the player to be recommended for selection, she preferred to use the average because it is more accurate and the average showed clearly the class that performed better. Excerpt B14 describes the above points.

Task 15. Initially Bella mentioned that 5C did better. She compared the scores of both classes without realizing the difference in the number of students involved in both the classes. However, later she calculated the mean and decided on the class that performed better based on this.

Table 4.55

Excerpt B15

R: Based on the graphs below, which class did better? S: 5C did better. R: Why? S: If I see the scoring, 5C is higher than 5D. If we look at the number of corrects, only two corrects and 9 corrects have the same number of students. The others, 5C has more students as compared to 5D. R: How about the students? S: Oh. The number of students is not equal. 5C is 36 students and 5D 1 is 21 students. I want to calculate the mean first. (Starts working on the mean calculation) R: Which class did better? S: 5D got higher average. So 5D did better R: You mentioned graph and then you said average. Which is the best way to determine the class that did better? S: 5C got 36 students, can say that each of the students in 5C get 5.5 corrects meanwhile 5D got 21 students so each the students in 5D gets 6.19 corrects. So, when we compare this, each student in 5D got better score compared to each student in 5C. So, students in 5D performed better. I would use the average because the graphs are misleading. The number of students in both classes are not the same.

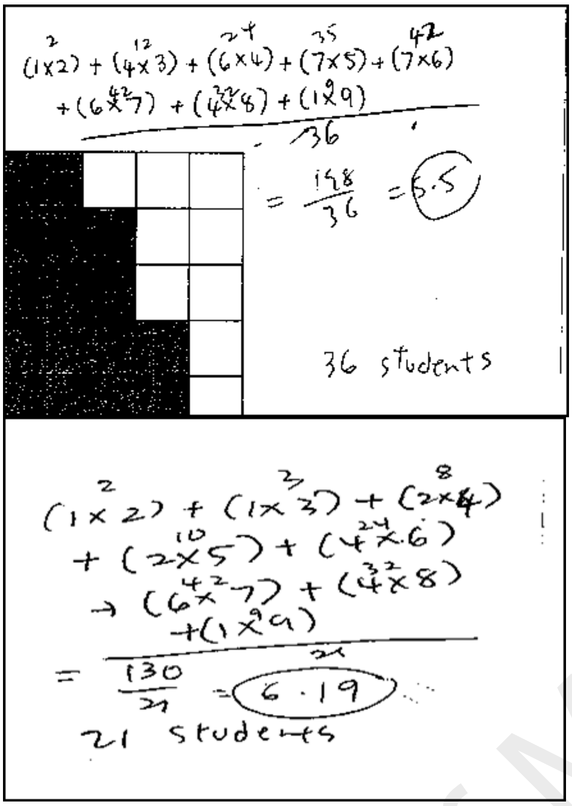


Figure 4.12. Entry B5

Initially Bella decided that 5C did better. She compared the graphs and looked at the students' scores. She claimed that 5C had higher scores compared to 5D. She said in her example that the number of students who got two corrects and nine corrects for both classes were the same. However, for the rest of the number of corrects, 5C had more students compared to 5D.

Bella only realised that the number of students were not equal after she was probed. She realised that 5C had 36 students and 5D had only 21 students. After when Bella mentioned that she needed to calculate the mean and started working on the calculation. Bella's mean calculation is shown in Entry B5.

Now she claimed that 5D did better because the class average is higher. She explained that each student in 5C got 5.5 corrects whereas each student in 5D got 6.19 corrects. She added when she compared this, 5D students got better scores than 5C. Thus, she claimed now 5D performed better. As for the two different ways that she mentioned she would use to choose the class that performed better, Bella decided that she would use the average instead of the graphs. She felt that the graphs can be misleading when the number of students in both classes are not the same. The above Excerpt B15 describes Bella's points on this.

Summary.

Table 4.56

Subject matter knowledge in making inference

Task	Statistical element	Conclusion
Task 12	Summarises equal sized numerical data using measures of central tendency	Bella summarised equal sized numerical data using the mean.
	Utilises the appropriate measure to compare and draw conclusion about two equal sized numerical data	Bella utilised the mean to compare and draw conclusion about two equal sized numerical data. In fact, Bella had mentioned that a one to one score comparison can be used too but she said that among these two methods the mean or the average is the best to be used to compare and select the player.
Task 13	Summarises unequal sized numerical data using measures of central tendency	Bella summarised unequal sized numerical data using the mean.
	Utilises the appropriate measure to compare and draw conclusion about two unequal sized numerical data	Bella said in order to compare and draw conclusions based on two unequal sized numerical data, the mean or the average is the only and best way to do so. She said that the one to one game score comparison cannot be done here because of two unequal sized data. Bella said that the mean is fair because it considered all the games the player had played.
Task 14	Summarises equal sized graphical data using measures of central tendency	Bella summarised equal sized graphical data using the average or the mean.
	Utilises the appropriate measure to compare and draw conclusion about equal sized graphical data	Bella had two different ways of identifying the player that she would recommend to be selected for the team. She said it can be based on the average score comparison or by looking at the graph. However, Bella preferred the average score comparison because she claimed it is more clear and accurate.
Task 15	Summarises unequal sized graphical data using measures of central tendency	Bella summarised two unequal sized graphical data using the mean.

	Utilises the appropriate measure to compare and draw conclusion about unequal sized graphical data	Bella initially compared the scores of students to decide on the class that performed better. However, after realizing that both the classes are of unequal size, she decided to use the average to compare and draw conclusion on the class that performed better. She said that due to the unequal sized data, the graphs can be misleading.

Level of subject matter knowledge of measures of central tendency.

Level of subject matter knowledge of measures of central tendency with reference to context. In Task 1, Bella mentioned that the average was found by adding up all the children from all the families and then divided the total with the number of families. She also showed the mean calculation in her example. Bella knew that for the context of the given data which was referring to the number of children, the ideal average has to be a whole number. However, she said that any division of two numbers in the mean calculation can result in a decimal average. She also said that this decimal average can be rounded off to a whole number to make it more realistic given to the context of the data. Therefore, Bella had presented the knowledge of the mean as the average (P) which is considered as appropriate (A).

In Task 2, Bella had utilised the idea of mode without the task explicitly stating so. She found the frequency for each shoe size and had identified the mode as 5 based on the highest frequency. Therefore, Bella's idea of the mode was found present (P) and correct (A).

Bella said that the mode can be a form of data representation. She explained that a lot people wore size 5. Therefore, Bella's knowledge of the mode as a form of data representation was found to be present (P) and appropriate (A).

However, Bella did not utilise the mode as the average. Instead Bella calculated the average as 5.85. She also said that she would not consider the mode as the average because the average that she calculated was 5.85 and when rounded off would be 6 and not 5. She also explained that according to the exact definition of the terms, the definition of the mode is not the same as the average. Bella had lacked in the knowledge that the mode can be used as a quick method in the reporting of an average.

Therefore, Bella's knowledge of the mode as the average was not only found to be not present (NP) but also inappropriate (I).

In Task 3, Bella identified the mode as 22. She took the mode as the highest number of comments. When the task was modified where the number of comments for student D was changed from 22 to 6, Bella identified the mode as 6 in which she still referred to the highest number of comments. Bella had identified the mode incorrectly. First, she had taken the highest number of comments to be the mode in which she believed to be highest frequency. Bella also lacked in the knowledge that the mode can be categorical. The mode is referred to be the category that carried the highest frequency. Therefore, Bella's idea of the mode was marked to be not present (NP) and the lack is considered problematic (I).

Bella mentioned that she would not utilise the mode to represent the number of comments made by the students. She said that the mode was found to be different than the other values in the data (for the case where the number of comments for student D was 22). Bella also mentioned that she would only utilise the mode to represent the data if the average that she calculated was around the same value as the mode (for the case where the number of comments for student D was 6). Bella lacked the knowledge that the mode can be utilised as a form of data representation. Therefore, Bella's knowledge of the mode as a form of data representation was found to be not present (NP) and the lack was considered problematic (I).

In Task 4, Bella mentioned that the average means the most. However, her later explanations revealed that her knowledge of the average is the mean in which she elaborated on the mean calculation in her example. Bella said that in the calculation of the average, all the data including the extreme data need to be considered. Bella also mentioned that she is not aware of any other ways in finding the average except for the mean calculation.

The average based on any data involving human population such as the Malaysian primary school students is normally based on the median because the median is definitely a more robust measure when dealing with data that contained extreme data. Therefore, Bella did not present the knowledge of the median as the average (NP). However, Bella's knowledge of the mean as the average is acceptable because an average can be based on any of the three measures of central tendency but her explanation involving the mean calculation that included the extreme data was found to be inappropriate therefore her justification was considered inappropriate (I). The appropriate justification would have been that the mean considered excluded the extreme data.

In Task 5, Bella found the median by rearranging all the weights from the smallest to the biggest. Bella added the weights of all the nine children and then divided the total with two in which she got 89.5. She started adding the weights one by one from the smallest weight in the arrangement until she came to a total of 85 which was at 19. Bella said that when she added the subsequent 19, the total exceeded 89.5. Therefore, she mentioned that the total 89.5 is somewhere in between the two 19s which she averaged these two numbers and found the median as 19.

Bella carried out a similar procedure to find the median for ten children after 43 kg was included in the data. She totalled up all the weights and divided with two. She got 111 in which she tried to locate the position for this value in the arrangement. She found that the position was somewhere between 19 and 24. She averaged these two numbers and got the median as 21.5. Bella's knowledge of the idea of the median was found to be not present (NP) because she had totally carried out an incorrect procedure to find the median which is considered to be problematic (I).

Bella said that the mean is not adequate to represent the weight of all ten children. She explained that this was because the weight of the tenth child was much heavier compared to the rest. Instead Bella said to consider the weight of the nine children and exclude the weight of the tenth child in the mean calculation. Bella did not discuss on the suitability of the other measures of central tendency in representing the weight of all ten children. Therefore, Bella's knowledge related to the robustness of the measures of central tendency was found to be not present (NP) because she lacked in the knowledge of the median being a more robust measure in the presence on an extreme data. However, Bella's justification towards the mean that was calculated after the extreme data was excluded was found to be appropriate (A).

In Task 6 and Task 7, Bella constructed data sets based on the given average value and the conditions imposed. In Task 6, Bella placed the prices randomly but around the average price. She utilised the backward mean calculation and got the total price as 189. She said that she placed the first random price and subtracted this from the total price. She repeated the same with the subsequent prices too until she came to the seventh price which gave the balance total price as zero. Bella said that she based her prices around the average price. Therefore, Bella's knowledge of constructing a data set based on the idea of average was present (P) and found to be appropriate (A).

In Task 7, Bella placed the prices in a similar way as what she did in Task 6. She calculated the total price. Then she placed the first random price in which she subtracted this from the total price. She repeated this with the subsequent prices too until the seventh price which gave a zero balance of total price. Bella mentioned that she considered the values that are close to the average price. Bella's knowledge of constructing a data set based on the idea of average for Task 7 was found to be present (P) and appropriate (A).

Table 4.57

Coding rubrics for determining level of Bella's subject matter knowledge of measures of central tendency with reference to context

Task	Subject Matter Knowledge of Measures of Central Tendency with reference to Context																															
	Mean as average				Mode as average				Median as average				Idea of mode				Idea of median				Mode as data representation				Robustness of measures				Idea of average			
	P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP	
	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I
Task 1	X																															
Task 2								X					X								X											
Task 3																X								X								
Task 4												X																				
Task 5																				X							X					
Task 6																													X			
Task 7																													X			
	1												**1**								**1**						**1**		**2**			

Legend: P = Present NP = Not Present A = Appropriate I = Inappropriate

Bella's Percentage of Subject Matter Knowledge of Measures of Central Tendency with Reference to Context

$$= \frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)} \text{ X } 100\% = \frac{\mathbf{6}}{\mathbf{11}} \times \mathbf{100\%} = 54.55\% \text{ (Medium)}$$

Level of subject matter knowledge of measures of central tendency in handling bias. In Task 8, Bella calculated the mean for all nine data. She added up all the nine data and then divided the total with nine. She got the mean as 7.18. Therefore, Bella not only identified that the mean can be obtained but also summarised the data using the mean by calculating the mean value. Therefore, Bella's knowledge of identifying and summarising the data using the mean was found to be present (P) and correct (A).

Bella said that the median can be found after the data is arranged in ascending order. She added up all the nine data and divided the total with two. She got 32.325. Bella said that she found the position for 32.325 using the cumulative values which she added from the lowest value in the arrangement, one by one until she reached to the position for 32.325. Bella found that 32.325 was positioned in between 6.2 and 6.3, so she averaged these two numbers and got the median as 6.25. Bella had actually performed incorrectly the procedure to find the median. Her median was also incorrect. Thus, Bella's knowledge of identifying and summarising the data using the median was found to be not present (NP) and considered incorrect (I).

Bella at first identified the mode as 15.3. However, she realized her mistake and changed her answer to 6.3. She said that 6.3 occurred thrice. Therefore, Bella's knowledge of identifying and summarising the data using the mode was found to be present (P) and correct (A).

Bella mentioned that the mean is the best measure to represent the actual weight. She said that the mean considered all the data. According to Bella, most of the data are around 6 to 7 and only one data is far from the rest which is 15.3, so she said that the mean that used all the data, represented the actual weight. Bella's knowledge related to the best representation measure is found to be present (P) because the mean can be used as the best representation measure. However, her explanation that the mean which considered all the data including the extreme as the best representation measure was considered inappropriate (I). The best representation mean would be the one calculated after excluding the extreme value or the outlier because this mean represents the data better.

Similarly, in Task 9, Bella calculated the mean for all five data. She added up all the five data and then divided the total with five. She got the mean as 76 800. Bella not only identified that the mean can be obtained but also summarised the data using the mean by calculating the mean value. Therefore, Bella's knowledge of identifying and summarising the data using the mean was found to be present (P) and correct (A).

Bella found the median as 48 000. However, Bella was not very consistent with the procedure to find the median. She got a value of 192 000 from the total of all the salaries which was divided by two. However, she neither arranged the data according to any particular order nor did she locate the position of the value using the cumulative total starting from the lowest value in the data like what she did in Task 8.

Instead, Bella started to locate the position for 192 000 from the original data. She added the salary one by one starting from the left side of the data until she found the position for 192 000 which she said was within the data value 48 000 itself. Therefore, she claimed that 48 000 as the median. Bella had actually performed incorrectly the procedure to find the median. Her median was also incorrect. Thus, Bella's knowledge of identifying and summarising the data using the median was found to be not present (NP) and considered incorrect (I).

Bella was also not consistent with the identification of the mode. In this task, she identified the mode as 180 000 which she said is the highest salary. Therefore, Bella's knowledge of identifying and summarising the data using the mode was found to be not present (NP) and incorrect (I).

Bella mentioned that the mean is the best measure to represent the actual salary because the mean considered all data. However, here Bella mentioned that the extreme salary will have to be tested first to see if this salary is accepted or rejected in the mean calculation. She said if the salary is accepted than the mean value which she calculated based on all five data would be the best representation mean but if rejected than she said the mean calculated based on the remaining four salaries would be the best representation mean. Bella's knowledge related to the best representation measure is found to be present (P) because the mean can be used as the best representation measure. Bella justified that the best representation mean would be based on a test to rule out any outliers. Her justification here is found to be appropriate too (A).

Table 4.58

Coding rubrics for determining level of Bella's subject matter knowledge of measures of central tendency in handling bias

Task	Subject Matter Knowledge of Measures of Central Tendency in Handling Bias															
	Identifies and summarises data using mean				Identifies and summarises data using median				Identifies and summarises data using mode				Best data representation measure			
	P		NP		P		NP		P		NP		P		NP	
	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I
8	X							X	X					X		
9	X							X				X	X			
	2								**1**				**1**			

Legend: P = Present NP = Not Present A = Appropriate I = Inappropriate

Bella's Percentage of Subject Matter Knowledge of Measures of Central Tendency in Handling Bias

$= \frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)}$ X 100% = $\frac{\mathbf{4}}{\mathbf{8}} \times \mathbf{100\%}$ = 50% (Medium)

Level of subject matter knowledge of measures of central tendency in problem solving. In Task 10, Bella carried out the backward mean calculation to obtain the total grade point average for College A based on two semesters as 6.4. She also found the total grade point average for College B based on three semesters as 11.4. Therefore, Bella's knowledge of performing the backward mean calculation to obtain the total from the mean is found to present (P) and correct (A).

Bella explained that the mean which was given for each college was for the number of semesters involved in that particular college. She also added the total grade points of both colleges and the total semesters of both colleges before she proceeded in the weighted mean calculation. Thus, Bella's knowledge related to the representative nature of the mean is found to be present (P) and appropriate (A).

Finally, Bella carried out the forward mean calculation to find the weighted mean. She divided the total grade points of both colleges with the total semesters of both colleges and found the weighted mean as 3.56. Therefore, Bella's knowledge of the forward mean calculation is found to present (P) and appropriate (A) despite the task not explicitly stating to find for the weighted mean.

In Task 11, Bella performed the backward mean calculation and obtained the total hours of TV watched per weekend for both rural and urban groups. She said that she multiplied the number of students in each group with the group average. Bella found the total hours of TV watched for rural group by multiplying 25 with 8. Similarly, for the urban group she multiplied 75 with 4. Thus, Bella's knowledge of performing backward mean calculation was found to be present (P) and appropriate (A).

Bella explained that she multiplied the average of each group with number of students in each group because the given average was for the group. She also mentioned that the totals of both the groups were added and the total number of students of both groups were found before she proceeded with further calculation because she wanted to find the overall average for all the students. Thus, Bella's knowledge related to the representative nature of the mean is found to be present (P) and appropriate (A).

Finally, Bella carried out the forward mean calculation to find the weighted mean. She divided the total number of hours of both groups with the total number of students and found the average as 5. Therefore, Bella's knowledge of the forward mean calculation is found to present (P) and appropriate (A) despite the task not explicitly stating to find for the weighted mean.

Table 4.59

Coding rubrics for determining level of Bella's subject matter knowledge of measures of central tendency in problem solving

Task	Subject Matter Knowledge of Measures of Central Tendency in Problem Solving											
	Backward mean calculation				Representative nature of the mean				Forward mean calculation			
	P		NP		P		NP		P		NP	
	A	I	A	I	A	I	A	I	A	I	A	I
10	X				X				X			
11	X				X				X			
	2				**2**				**2**			

Legend: P = Present NP = Not Present A = Appropriate I = Inappropriate

Bella's Percentage of Subject Matter Knowledge of Measures of Central Tendency in Problem Solving

$$= \frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)} \text{ X } 100\% = \frac{\mathbf{6}}{\mathbf{6}} \times \mathbf{100}\% = 100\% \text{ (High)}$$

Level of subject matter knowledge of measures of central tendency in making inference. In Task 12, Bella calculated the mean scores for both the players by adding all the scores and divided the total with the six set of games that they played. She found that the mean score for Ramu was 19.67 and the mean score for Beng was 21.83. Therefore, Bella's knowledge of summarising equal sized numerical data using the measures of central tendency in this case the mean without the task explicitly stating to do so was found to be present (P) and appropriate (A).

Bella compared the mean scores of both the players and decided on the player that she would recommend to be selected for the team. Since Beng's average was higher than Ramu's, she said that Beng would be recommended. Bella also mentioned that the player can also be selected by comparing each set of games. The player that won the most sets would be selected. However, Bella mentioned that the average comparison is the best way to compare and decide on the player to be selected because the average considers all the scores for each player. Thus, Bella's knowledge of utilising the appropriate measure of central tendency to compare and draw conclusion on two equal sized numerical data was found to be present (P) and appropriate (A)

In Task 13, Bella summarised the data using the mean or the average. She calculated the average by adding all the scores for each player and then divided the total scores with the number of game sets that they played. She said for Ali she divided his total scores with eight because he played eight game sets. Whereas for Khan, she divided his total scores with six because he played six game sets. Bella found that Ali's average was 22.25 while Khan's average was 23. Therefore, Bella's knowledge of summarising unequal sized numerical data using the measures of central tendency in this case the mean without the task explicitly stating to do so was found to be present (P) and correct (A).

Bella compared the average scores of both the players and decided to recommend Khan because his average score was higher than Ali's. Bella said that the average was a fair measure because all the game sets that the player played were taken count. Bella said that since the number of game sets played by the players were different, therefore the one to one game comparison cannot be done here. She added that only the average comparison can be done.

Thus, Bella's knowledge of utilising the appropriate measure of central tendency in this case the mean to compare and draw conclusion on two unequal sized numerical data was found to be present (P) because the mean is the most suitable measure to be used to compare and draw conclusions based on unequal sized data. Bella's knowledge of utilising the mean here was considered to be appropriate too (A).

In Task 14, Bella summarised the data given using the mean or the average. She multiplied the number of students with the number of corrects and added the products. Then she divided the total with the total number of students. Bella calculated the average for 5A as 4.89 and for 5B as 4.22. Therefore, Bella's knowledge of summarising equal sized graphical data using the measures of central tendency in this case the mean without the task explicitly stating to do so was found to be present (P) and her knowledge of calculating the mean was found to be appropriate (A).

Bella compared the class average of both the classes and found that 5A had a higher average than 5B. Therefore, she said that 5A performed better than 5B. Bella also mentioned that by comparing the graphs of both classes, the class that performed better can be determined too. She said that 5A students managed to get more corrects as compared to 5B. However, she said that the average comparison is more accurate because the values indicated clearly the class that performed better. Thus, Bella's knowledge of utilising the appropriate measure in this case the mean to compare and draw conclusion on two equal sized graphical data was found to be present (P). She also had calculated the average or the mean correctly and explained that the class that had higher average performed better, so her justification here is considered appropriate (A).

In Task 15, Bella summarised the data given using the mean or the average. She multiplied the number of students with the number of corrects and added the products. Then she divided the total with the total number of students. Bella calculated the average for 5C as 5.5 and for 5D as 6.19. Therefore, Bella's knowledge of summarising unequal sized graphical data using the measures of central tendency in this case the mean without the task explicitly stating to do so was found to be present (P) and her knowledge of calculating the mean was found to be appropriate (A).

Bella compared the scores for both classes and said that 5C performed better. She did not realize that both classes had unequal number of students. After realising that the number of students were not equal, Bella calculated the mean. Based on the mean or the average comparison, Bella said that 5D did better because the average was higher than 5C. Bella said that since the number of students were unequal, the graphs are misleading cannot be compared. She said that she preferred to use the average. Thus, Bella's knowledge of utilising the appropriate measure in this case the mean to compare and draw conclusion on two unequal sized graphical data was found to be present (P). She also had calculated the average or the mean correctly and explained that the class that had higher average performed better, so her justification here is considered appropriate (A).

Table 4.60

Coding rubrics for determining level of Bella's subject matter knowledge of measures of central tendency in making inference

Task	Subject Matter Knowledge of Measures of Central Tendency in Making Inference																															
	Summarises two equal sized numerical data using measures of central tendency				Utilises the appropriate measure to compare and draw conclusion about two equal sized numerical data				Summarises two unequal sized numerical data using measures of central tendency				Utilises the appropriate measure to compare and draw conclusion about two unequal sized numerical data				Summarises two equal sized graphical data using measures of central tendency				Utilises the appropriate measure to compare and draw conclusion about two equal sized graphical data				Summarises two unequal sized graphical data using measures of central tendency				Utilises the appropriate measure to compare and draw conclusion about two unequal sized graphical data			
	P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP	
	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I
Task 12	X				X																											
Task 13									X				X																			
Task 14																	X				X											
Task 15																									X				X			
	1				**1**				**1**				**1**				**1**				**1**				**1**				**1**			

Legend: P = Present NP = Not Present A = Appropriate I = Inappropriate

Bella's Percentage of Subject Matter Knowledge of Measures of Central Tendency in Making Inference

$$= \frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)} \text{ X } 100\% = \frac{\mathbf{8}}{\mathbf{8}} \times \mathbf{100\%} = 100\% \text{ (High)}$$

Bella's mean percentage of subject matter knowledge for all four constructs $\frac{\mathbf{54.55+55+100+100}}{\mathbf{4}} = \mathbf{76.14}\%$. Therefore, Bella's level of subject matter knowledge of measures of central tendency is medium.

Case Study Three: Harry

Harry was 24 years 5 months old when he was interviewed. At the point of data collection, he was pursuing a 4-year Bachelor of Science with Education (B.Sc. Ed.) program at a public university. He majored and minored in mathematics and chemistry respectively. Harry obtained an A1 in Mathematics and an A2 in Additional Mathematics in his SPM examination (equivalent to O level examination). He also scored an A- in Mathematics T in his STPM examination.

At the time of data collection, Harry was in the final semester of fourth year studies. He attained 3.18 in the Cumulative Grade Point Average (CGPA) for his first three years of studies in the public university. He does not have any teaching experience prior to this interview. The following sections discuss the findings of Harry's subject matter knowledge of measures of central tendency in each of the four different constructs, namely with reference to context, in handling bias, in problem solving, and in making inference that were emerged from the clinical interview.

Types of subject matter knowledge.

Subject matter knowledge of measures of central tendency with reference to context.

Task 1. Harry's knowledge of the mean is average. He mentioned that non-divisible numbers can result in a decimal average.

Table 4.61

Excerpt H1

R: How do you think the average was obtained?
S: The average is the sum of children divided by the number of families that were taken in. For example, there are five families, different families will have different number of children. After add up and divide by the number of families, then average 2.3 children is obtained.
R: Can you show and explain using an example?
S: For example, we take five families and each family has 3, 8, 4, 2, 5 children. So, to find the average, we just add up 3, 8, 4, 2, 5 divided by five families. We get 22 divide by 5. We get 4.4 children per family. This is the average.
R: How can the average be 2.3, and not a counting number like 1, 2, 3, or 4?

S: {Looked a bit unsure}. Maybe the total number children that we get for the families is not divisible by the total number of families that we considered. For example, if total we get 20 divided by 5, so will be a whole number so each family will have 4 children. So maybe the total number here is not divisible with the denominator so we get a decimal number. 2.3 basically is not logic because children, we count as per person. People, we count as one by one.

Harry mentioned that the average is obtained from the sum of children divided with the number of families taken into account. He elaborated that if there are five families, these families will have different number of children. He said that the average 2.3 is obtained after the number of children for these families are added up and then divided with the number of families.

Harry's explanation on the example also describe further on this. He mentioned that for example if there are five families and each family has 3, 8, 4, 2, 5 children. In order to find the average, he added up 3, 8, 4, 2 and 5 then divided with five families. Harry divided 22 with 5 and got the average as 4.4 children per family.

However, Harry looked rather unsure when he was asked to explain on how the average can be 2.3 and not a counting number like 1, 2, 3, or 4. He explained this might be due to the total number of children for the families is not divisible by the total number of families considered. He elaborated that if two divisible numbers like 20 and 5 are divided, then the average will be a whole number. However, he said if the total number of children is not divisible with the denominator, then a decimal average is possible. Harry also mentioned that 2.3 for children is not logic in which he said that children has to be counted as per person or one. The above Excerpt H1 illustrates Harry's point on this.

Task 2. Harry utilised the idea of mode in real life situation without the task explicitly stating to. However, he did not utilise this concept as a form of data representation. Harry mentioned that in certain situation, the mode can be used as the average.

Table 4.62

Excerpt H2

R: How would you place the order for the shoe sizes in your shop? S: Arrange the sizes first from 4, 5, 6, 7, 8, 9. then count one by one, for example size 4 got how many people who purchased this shoe size. And so, for the rest. Then I order based on the shoe size that is purchased the most, I order the most. The rest follows according to the volume of purchase. R: If you want to choose one female shoe size, which shoe size would you choose? S: I would choose size 5. R: Why? S: Because females that are wearing size 5 is the most among all the other shoe sizes. So that means the demand for size 5 is the highest. So, we order size 5 more than the others. R: Can the chosen shoe size represent the female shoe size in your shop? S: Cannot R: Why? S: Because the chosen shoe size is 5 but there are other females that also wear sizes that are smaller than or bigger than this size. So, we cannot straight away represent the female shoe size in our shop as size 5. R: What would you use as the average shoe size? S: The average shoe size is basically for us to use when ordering shoes. For the shop or other company. From the data, we know that size 5 is the hottest size because many wear this size, so we use the average for us to order the stock. So, when customer come, most of them are size 5 so we got stock to sell to them. R: How do you find the average here? S: Average here I straight away look at the highest frequency 12. So, the average is 5. R: In regards to measures of central tendency, what is average is referring to? S: Here it is referring to the mode. R: How? S: Here we do not have other choices because size 5 is worn by most of the females with the highest frequency. So, this is the only way where we choose the mode as the average. We calculate the mean 4 times 5, 5 times 12, and so… and then divided by the total frequency that does not indicate any information that most of the females wears size 5. R: What would you use as the average shoe size? S: Size 5 R: Can you explain why? S: Because from the information, we can see that most of the females are wearing size 5, so this size 5 has a frequency of 12. It indicates that most of the females, their size is 5 so we choose this as the average.

Harry mentioned that he would place the order for the shoe sizes in his shop by first arranging the shoe sizes from the smallest to the biggest as the following 4, 5, 6, 7, 8, and 9. Then he said that he would count the frequency for each shoe size. He mentioned that he would order the most for the shoe size that is purchased the most and the rest of the shoe sizes will follow according to the volume of purchase.

Harry mentioned that he would choose size five if he was given a choice to choose one female shoe size. His decision was based on that size 5 was worn by most females as compared to the other sizes. Therefore, the demand for size 5 is the highest and more orders will be placed for this size compared to the others. However, Harry mentioned that the chosen shoe size 5 cannot represent the shoe sizes in his shop. He mentioned that there are also female customers who wear sizes that are either bigger or smaller than this size. So, he said that size 5 cannot straight away represent the female shoe size in his shop.

Harry mentioned that the average shoe size can be used for ordering the stock for the shoes either for his shop or other companies. When Harry was asked on how he would find the average, he mentioned that he would look at the highest frequency 12, therefore the average is 5.

Harry said that in relation to measures of central tendency the average can be referred to the mode. He explained that in this situation, there are no other choices because size 5 was worn by most of the female customers which was indicated by its highest frequency. Therefore, he claimed that here the mode can be used as the average. He added that if the mean is calculated, it does not indicate any information on the size 5 which is worn by most of his female customers. The above Excerpt H2 describes Harry's point on this.

Task 3. Harry obtained the mode based on the category that carried the highest frequency. He expressed that the mode cannot be used as a form of data representation because it represents just a single category. He mentioned that the mean is more suitable because it considers all the data in the set.

Table 4.63

Excerpt H3

R: What is the mode?
S: The mode is D looking at 22.
R: How did you obtain this value as the mode?
S: Straight away look at the highest frequency from the data and you get the mode.
R: If you were to represent one value for the number of comments made by the students on that day, would you use this value?
S: No
R: Why?
S: Because from the data, some of students never gave any comments and some gave a lot of comments until 22. So, we cannot straight away take the mode to represent the number of comments for all the students. But I can use the mean to represent.
R: How?
S: Just add up the number of comments then divide by the number of students. The mean is 4.6 and round off to 5 because comments should not be in decimal number.
R: If I change the value for D here from 22 to 6. What is the mode now?
S: The mode is still D.
R: How did you obtain the mode?
S: Because of the highest frequency like just now.
R: If you were to represent one value for the number of comments made by the students on that day, would you use this value?
S: Still no
R: Why?
S: Because the mode is based on the highest frequency but it cannot represent the rest of the information. There is still zero comment from student A. I will still use the mean.
R: Why?
S: Because to balance up the zero comments from student A and some of the students who commented a lot. We just take the average. So, the mean will represent the number of comments made by the class on that day. Because we must understand that each number of comments here is not representing the whole class but representing a particular student therefore the mode to represent the number of comments of the class is not suitable.

Harry identified the mode as student D based on the highest number of comments which is 22. He said that the highest number of comments referred to the highest frequency for the data. Harry mentioned that he would not use the mode to represent the number of comments because from the data, he noticed that some students never made any comments and some made a lot of comments until 22.

Therefore, he mentioned that the mode cannot be taken to represent the number of comments made by all the students. Instead Harry said that he would use the mean to represent the number of comments. He explained that he would just add all the number of comments and then divide with the number of students. He got the mean as 4.6 in which he rounded off to 5 because the number of comments should not be in decimal.

The task was later modified where the number of comments for student D was changed from 22 to 6. In this situation too, Harry identified the mode as student D in which he still based on the highest number of comments that he claimed as the highest frequency. He mentioned that he still would not use the mode to represent the number of comments made by the students on that day because the mode is based on the highest frequency and cannot represent the other information. He said that there still exist the zero comment from student A

Instead, Harry preferred to use the mean because he said that the mean would balance up the zero comments from student A and some of the students who commented a lot. He mentioned that the mean can represent the number of comments made by the students on that day. He explained that each number of comments is not representing the whole class but a particular student. In such he said that the mode is not suitable to represent the number of comments. This is described in the above Excerpt H3.

Task 4. Harry related the average to the mean. He also said that the average can be based on the mode if the sample size was big enough. However, He claimed that the median is not suitable to be used as the average.

Table 4.64

Excerpt H4

R: What does average mean in this sentence?
S: Average here means that most of them watch 3 hours of TV per day.
R: Can you explain further?
S: The total numbers of hours obtained from the sample divide by the total number of students in the sample. Here we will get the average.
R: How do think they got this average (3 hours of TV per day)?
S: They found the mean.
R: Can you explain further using an example?
S: For example, we take 6 students. A small sample. In the sample of 6 students we add 3, 4, 1, 9, 12, 5 hours. After divide, we get 5.67 hours. That means for this group, the average per student is 5 hours. For me this is the idea to calculate average.
R: There is an information saying that a small group of students watched more than 12 hours of TV per day, so how do you think they deal with information here?
S: This information totally will affect the average because some of them maybe watch just 1 hour and after we add 12 hours or more, it will bring the 1-hour value up. For example, I only watch 1 hour and after I count this in maybe my average becomes 3 hours which is more 2 hours than what I watch.
R: So, what do you think they would have done with that information?
S: It will affect the average. Maybe the average may not be accurate. The value might deviate from the actual.
R: Is there any other ways that they would have find the average?
S: We need to see how big the sample is. If the sample is more maybe, we can use mode.
R: How? Please explain
S: Because when the sample is big, we can use mode for example 3 hours is the mode in a sample of 100. Those who watch 12 hours maybe is just only below 10. So, the mode can represent that most of the students in Malaysia watch 3 hours of TV per day. So yes, the mode can be used to represent if the sample is big enough.
R: Are you saying that you can use the mode as the average?
S: Yes
R: What about the median?
S: Median {looked unsure} median I think not suitable.
R: Why?
S: Because the median is the centre of the data, for example, 1 to 24 hours, median is 12 hours. So, 12 hours cannot represent the rest. We know the graph is like this (points at the graph in his task sheet). If like this than most of the students watched 12 hours but from the information is 3 hours. So, median cannot be used.

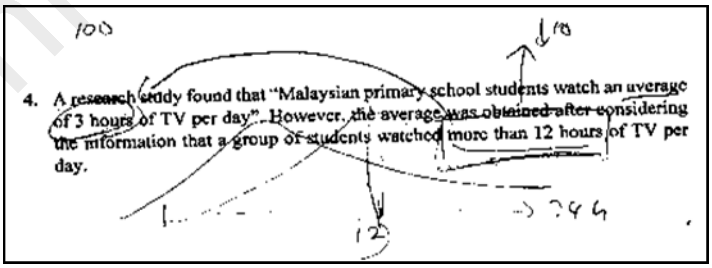


Figure 4.13. Entry H1

Harry mentioned that the average can mean that most of the students watched 3 hours of TV per day. He explained that the average can be obtained by dividing the total number of hours students watched TV per day divided by the total number of students from a sample. Harry mentioned clearly that the average here was obtained from the mean. He elaborated this further in his example where he mentioned that for a sample of 6 students where each of them watched 3, 4, 1, 9, 12, and 5 hours of TV per day, the number of hours is added and then divided with the total number of students. He calculated the average as 5.67 in which later he rounded off the figure to 5 hours of TV per day. Harry added that this was his idea to find the average.

Harry mentioned that the information on a group of students watched more than 12 hours of TV per day will affect the average. He elaborated that some might have watched 1 hour and some 12 hours or more, therefore the information on 12 hours or more will bring the 1-hour value up. He said this will also affect the average which will also be raised. He added that the average might not be accurate and might be deviated from the actual.

On the other hand, Harry said that if the sample size is big than the mode can be considered as the average. He elaborated this in an example where for a sample of 100, the mode is 3 and less than 10 students watched more than 12 hours. Therefore, according to Harry, the mode can represent that most of the students watched 3 hours of TV per day. He claimed that in such situation where the sample size is big enough than the mode can be used as the average.

However, Harry was unsure if the median can be considered. He claimed that it is not suitable because the median refers to the centre of the data. He explained in his example that for 1 to 24 hours, the median here is 12 hours and this median cannot represent the rest of the data. He pointed this to a graph sketch that he made in the task sheet as shown in Entry H1. According to him, if the median is considered than most of the students watched 12 hours when it should be 3 hours. Thus, for him the median cannot be considered as the average. The above Excerpt H4 illustrates Harry's points on this.

Task 5. Harry obtained the median by arranging first the data from the smallest to the biggest value. He elaborated that if the data is odd numbered, the centre or the middle data in the arrangement is the median and if the data is even numbered, the median is obtained from the two middle values added and then divided by two. Harry mentioned that the median is a better measure to represent the data as compared to the mean because the median does not take count the outlier 43 kg.

Table 4.65

Excerpt H5

R: Which is the weight for the median child?
S: The weight for the median child is 18.5
R: How did you find the median?
S: 18 plus 19 divide by 2
R: Can you explain further?
S: We arrange the numbers from smallest to biggest. Then we try to count how many total values. If it is 8 an even number, we add the fourth and the fifth number and divide by two. If the number is odd number, we straight away take the middle value in the arrangement. [Looked at the working and realized a mistake in the median] Sorry a mistake. Ok the median will be 19. 19 is the median because just now I missed out one value in the arrangement which is 26.
R: Can you explain again?
S: From the information, first we try to arrange the weight of the children from the smallest to the biggest. Then if the total number of data that we get is odd number, we straight away determine the centre/middle. If it is even, then we need to plus the two numbers in the centre and divide by two. For this question, it is odd so straight away find the centre or the middle value which is four from the left and four from the right and we find that the fifth value the centre which is 19.
R: Which is the median if we include another child who weighs 43 kg?
S: If we include the child 43 kg. The median is still 19 kg because when we add 19 plus 19 divide by 2 is 19.
R: Is it adequate to use mean to represent the weight of the 10 children?
S: [Thinks for a while] let me try first. (Calculates the mean) Not suitable.
R: Why?
S: Because from this data, there is a child that weighs 43 kg. So, 43 kg is an outlier in this data which is far away from the rest. If we want to let this sample to be adequately represented than mean is not really ok. Because of this person (43 kg). If we add another person which the weight is almost closer to 26 or 27. Then we can use the mean to represent the weight of the children because the difference between each weight is small and not much.
R: How will you represent the weight of the 10 children?
S: Mmm… I think [paused a while]. The median is the suitable measure.
R: Why?
S: Because the median obtained represents the data. We can use the median to represent the children because median is the centre weight among these 10 children. If mean it will affect in the sense when we compare the mean with the smallest and the biggest data, we find that the difference is big but median the difference is not much. Median does not take count 43 kg the outlier.

Harry mentioned that the median is 18.5. He found this value after he added 18 with 19 and then he divided with two. He explained that he first arranged the weights from the smallest to the biggest value. Then he found the total number of data. He mentioned that in this situation, the total was 8 which was even numbered. Therefore, he said that the median is obtained after the fourth and the fifth data were added and then divided with two. He further explained that if the data is odd numbered then the middle data in the arrangement will be the median.

However, as Harry was explaining his answer, he realized that he had made a mistake. He apologized and quickly changed the answer for the median as 19. He said that he had left out one data in the arrangement which is 26. He explained again that since now the data is odd numbered the median is the fifth data which is in the centre/middle with four data on both sides of the arrangement.

Harry mentioned that the median is still 19 after 43 kg is included into the data. He obtained this by adding the two-middle data which are 19 and 19 then divided the total with two. Harry was unsure if the mean can be used to represent the weight of the 10 children. He took some time to think and then calculated the mean. After which he concluded not suitable.

He explained that this is because of the weight 43 kg which is an outlier far away from the rest of the data. Therefore, he felt the mean is not adequate to represent the weights of all ten children. He elaborated that this situation would have been different if 26 or 27 kg was included instead of 43 kg. In such case the mean would have been suitable to represent the weight of the children because the difference between each weight in the data is not much and small.

Harry paused and took some time to think on how he would represent the weight of the 10 children. He claimed that the median is a suitable measure because the median that he got does represent the data. Moreover, he said that the median is the centre weight for these 10 children. The difference between the median and the data is not much because the median does not take into account the outlier 43 kg. Unlike the mean, which he said the difference between the mean with the smallest and the biggest data is big. The above Excerpt H5 describes this.

Task 6 and task 7. Harry utilised the backward mean calculation to find the total price for all seven bags of crisps in both the tasks. The total was used to guide his data construction.

Table 4.66

Excerpt H6

R: How did you place the price on each bag of crisps?
S: I straight away put 27. All 27.
R: Why?
S: Because the average is 27, then 27 times 7 and divide by 7, we get the same value as the average which is 27.
R: How did you base these values?
S: Because the average is 27. So, I straight away put 27. 27 times 7 and divide back with 7, we get the average back as 27. If RM 27 is the average and since we are allowed to put any prices to get that average, I straight away take it as 27.

In Task 6, Harry placed the prices for all the bags of crisps as 27. He said that he straight away put 27 and all 27. He utilised the backward mean calculation to explain his decision. He explained that since the average is 27 and when the average 27 is multiplied with 7 and divided again with 7, the same value as the average 27 is obtained. Therefore, he straight away placed the prices as RM 27 because the task did not restrict him from doing this.

Table 4.67

Excerpt H7

R: How did you place the price of the bags here?
S: Since that none of the bags can cost 27, so we know that the total number of bags is 7. So, *x* divides by 7 equals to 27. When bring the 7 up and times 27, we get the total as 189. So, from all the different prices that we are going to place, the total must come up to 189 regardless what prices that we place. All must add to 189.
R: So how did you base all the individual values here?
S: I just simply choose any number like this as 20, this one as 29 and then after 189 minus 20 minus 29 and then get the value and then divide by 5. Each other values will be 28. But this is just

one possibility, there can be other numbers also. The thing that we need to be sure is that to make sure that all the prices that were placed must sum up to 189. We can put this as RM 1 but after total up all the seven prices must be adding up to 189. Since the task did not fix any condition on what price to be placed. Any value is possible.

On the other hand, for Task 7, Harry performed first the backward mean calculation to obtain the total of all prices from the average price. He said that none of the bags can cost 27 and that there are all together a total of seven bags. He elaborated that when x is divided with 7 will give the average price 27. Therefore 7 multiplied with 27 will give the total of all prices as 189. He mentioned that from all the different prices that he will place, the total of all the prices has to be 189.

Harry explained that he simply chose any two numbers such as 20 and 29. These two numbers were subtracted from the total price 189. The balance that he obtained, he divided with 5 which gave him 28. This value was placed for all the other remaining prices. However, he mentioned that this is just one possible combination. There are also other combinations based on other price values which can go as low as RM 1 because the task did not pose any restriction towards the price. Any value is possible. However, Harry did mention to make sure that the sum of all the prices placed add up to 189. The following Excerpt H7 illustrates Harry's point on this.

Summary.

Table 4.68

Subject matter knowledge with reference to context

Task	Statistical element	Conclusion
Task 1	Mean as average	Harry's knowledge of the average is synonymous to the mean. He elaborated on the mean calculation in his explanation on how the average would have been obtained and also in his example. Harry mentioned that the decimal average given to the context of the data is not logic but he said when non-divisible numbers are involved in a division, a decimal number is possible.
Task 2	Idea of mode	Harry utilised the idea of mode without the task explicitly stating to do so. He counted the frequencies for each shoe size and said that he would place the most purchase for the mode size.

	Mode as data representation	Harry said that the mode cannot be used to represent the shoe sizes in his shop. He explained that the mode size is just one size and there are also other shoe sizes to be considered in the shop.
	Mode as average	Harry utilised the mode as the average. He said that the average is used to place the order and since the mode size is purchased the most so the mode can be used as the average.
Task 3	Idea of mode	Harry identified the mode by looking at the highest number of comments in which he said indicated the highest frequency. Harry's knowledge of the mode here is categorical.
	Mode as data representation	Harry said that the mode cannot be used to represent the number of comments because there were also other students who made other number of comments. He said that the mode is based only on the highest frequency and not the rest of the comments. Instead Harry said that the mean is better to be used to represent the data. According to him, the mean balances up all the comments made by the class and not like the mode that represents one particular student.
Task 4	Median as average	Harry mentioned that the average means the most. However, in his explanation on how the average was obtained, he mentioned the mean and elaborated the mean calculation in his example. Harry said that any extreme data would affect the average by deviating the value from the actual. He said also that the mode can be another way to find the average if the sample is big enough. He explained that the extreme data here would not have an influence on the average if it is based on the mode. However, Harry said the average here cannot be based on the median because it is the centre of the data.
Task 5	Idea of median	Harry explained that to find the median, he would first arrange the data in ascending order and then find the middle number as the median if the data is odd numbered. If the data is even numbered, he said that he would add the two middle numbers and then divide the total with two.
	Robustness of measures	Harry mentioned that the mean is not adequate in representing the weight of all ten children because the mean for all ten children carries quite a big difference compared to the smallest and biggest data. Instead he said that he would use the median because the median represents the data and the difference between the median with the smallest and the biggest data is not much. He also said that the median does not take count the outlier.
Task 6	Idea of average	Harry placed all the values for the seven bags of crisps as the average price. He utilised the backward mean calculation in which he times the average price with the total number of bags and then divided it back with the total number of bags. He got the value same as the

		average price. Therefore, he said that he straight away took that value as the values for all seven bags of crisps.
Task 7	Idea of average	Harry utilised the backward mean calculation to find the total for all seven bags of crisps. He said that he would chose randomly any values and placed them. The values can be as small as RM 1. He said that he only made sure that all the seven values that he placed randomly adds up to the total price value.

Subject matter knowledge of measures of central tendency in handling bias.

Task 8 and task 9. Harry identified the measures of central tendency that can be obtained based on the data. In relation to the best measure of central tendency, Harry mentioned that the measure that had smaller difference with the lowest and the highest data would be the best measure to represent the data. However, Harry's justification towards the best measure in both tasks had different approaches when dealing with an extreme data. Harry excluded the extreme data when he considered the median but included the extreme data when he considered the mean as the best measure to represent the data.

Table 4.69

Excerpt H8

R: What are the measures of central tendency that you can obtain based on the above data?
S: Based on the above, actually all the three measures of central tendency can be obtained. Median, mean and mode
R: Can you explain to me, how you can get the mean, mode and median?
S: Median, we need to arrange again from the smallest value to the biggest value. 6.0, 6.0, 6.1, 6.2, 6.15, 6.3, 6.3, 6.3, 15.3. To find the median, we need to find from the arrangement which is the value that is the middle of the data arrangement. So, the median is the fifth data. 6.15 is the median.
R: What about the mean?
S: Mean, we need to add all the nine values that the students get and then divide by nine. (Calculates the mean). So, the mean is 7.18.
R: How about the mode?
S: For the mode we need to, arrange the number of values here that we got, for example 6.0 we need to find how many 6.0s there, we got 2. So, the frequency is 2. For 6.3 is 3. 6.1 is 1, 6.2 is 1, 6.15 is 1, 15.3 is 1. The mode is 6.3. Which there are 3 students.
R: Which is the best measure of central tendency to represent the actual weight of this object?
S: From here, I will choose median.
R: Why?
S: If we don't consider 15.3 because this value is totally out. We consider 6.0 until 6.3, median 6.15 is the most balanced one because the difference between 6.15 and 6.0 is 0.15 and then after eliminating 15.3 so the highest value now is 6.3. So, the difference between 6.15 and 6.3 is 0.15 also. So, we can see a balance between the lowest and the highest value after eliminating 15.3. So, I think median is the best measure. But if we choose mode after we eliminate 15.3, 6.3 is the

highest. So, the difference between 6.0 and 6.3 is 0.3 which just now median is 0.15 only. So, I think the median is the most suitable.
R: How about the mean?
S: No, I would not choose the mean because the mean that we get is already affected by 15.3. So, 15.3 has been divided and been shared by the rest of the values. That is why the mean that we get is 7.18. But from the data, the values are around 6 point something only. So, 7.18 is totally different and far away from the other values in the data. I would not choose the mean but I would choose the median.
R: Why would you choose the median?
S: I choose median because from the data, I can see clearly that median is the centre between after I ignore 15.3. Maybe 15.3 is an error. 6.0 is the lowest and 6.3 is the highest between the rests. So, when we choose the median, I can see that the difference between the median value 6.15 and the lowest value 6.0 is 0.15 and the same for the highest value 6.3 and the median value 6.15 is also 0.15. So, there is balance between the lowest and the highest. So, I choose median.

Harry mentioned that the mean, the median and the mode can be obtained from the data given in Task 8. He explained that in order to find the median, the values need to be arranged from the smallest to the biggest value. Harry elaborated that the order of the arrangement is as such 6.0, 6.0, 6.1, 6.2, 6.15, 6.3, 6.3, 6.3, and 15.3. He said that the median is the value in the middle of the arrangement which is the fifth data. He identified the median as 6.15.

According to Harry, the mean is calculated by adding all the nine values and then divide the total with nine. He calculated the mean as 7.18. Harry mentioned that the mode is found by first arranging the values with the respective frequencies. He said that the mode is 6.3 which had three students.

Harry mentioned that he would choose the median to represent the actual weight of this object. He said that he would not consider 15.3 because this value is totally out. He elaborated that he would consider the values from 6.0 until 6.3. Harry believed that the median 6.15 is the most balanced value because the difference between 6.15 and 6.0 is 0.15. He further added that the difference between 6.15 and the highest value in the data which is 6.3 after ignoring 15.3 is also 0.15. Therefore, Harry said that he could see a balance between the median with the lowest and the highest value in the data. This was the reason that he chose the median as the best measure to represent the actual weight.

Harry also explained that the mode is not suitable because the mode itself which is 6.3 is the highest value in the data after 15.3 is eliminated. On the other hand, the difference between the mode and the lowest value in the data is 0.3 as compared to the median which is only 0.15. Thus, Harry said that the median is still the most suitable measure.

On the other hand, Harry said that he would not choose the mean because the mean that he got is already affected by 15.3. He explained that 15.3 was shared with the rest of the values in the mean calculation which resulted in the mean value as 7.18. 7.18 is found to be totally different and far from the other data values which are around 6 point something. The above Excerpt H8 illustrates Harry's points on this.

Table 4.70

Excerpt H9

R: What are the measures of central tendency that you can obtain based on the above data?
S: Mean and median
R: What is the mean?
S: Mean is adding up 54 000, 42 000, 60 000, 48 000 and 180 000 then divided by 5. 76 800.
R: What about the median?
S: Median, I would arrange from smallest to the biggest. 42 000 is the smallest and the biggest is 180 000. So, median is 54 000 the middle number.
R: Is there any other ways that you can find the median?
S: We still need to arrange then 5 divide by 2 is 2.5 after round off is 3. So, the median is in the third place.
R: How about the mode?
S: No mode. Here totally there is no mode because every value exists as one value by itself.
R: Which is the best measure of central tendency to represent the actual salary of the company?
S: To represent the actual salary of the company, here considering all the given data I will choose mean.
R: Why?
S: If I choose median 54 000, cannot represent the actual salary of the company because there is a big value180 000. So, I will choose the mean. Because if we put this value here (pointing at the mean value). This value is situated here between 60 000 and 180 000. This means that here (42 000) and here (mean) the difference is 30 000 plus but here (mean) and here (180 000) is 90 000 or 100 000. But if we choose the median, it is much lower compared to this one (mean). That means the difference 180 000 and the median is bigger compare to choosing mean. So, to represent the actual salary of the company, I think we chose mean. {Looked unsure} [Mentioned that it is hard to explain]. Choose the one that has lesser difference. So, mean has lesser difference with the lowest and the highest value in the data as compared to the median. So that is why mean is chosen as the best measure to represent the actual salary.

Whereas in Task 9, Harry mentioned that only the mean and the median can be obtained based on the given data. He said that the mean is 76 800 which was calculated after all the five data were totalled up and then divided with 5. Harry mentioned that the median is 54 000. He obtained the median from the middle number after all the data was arranged from the smallest to the biggest value. Harry also mentioned that there is another way to find the median. He said that in this way, the data still needed to be arranged from the smallest to the biggest first but the total number of data 5 is divided with 2. This gave him 2.5 in which he rounded off to 3. Using 3, Harry mentioned that the median would be in the third place of arrangement. However, Harry said that there is no mode in this data because all the values exist only once.

In regards to the best measure to represent the actual salary of the company, Harry mentioned that he would choose the mean. He explained that the median cannot represent the data because there is one extreme salary which is 180 000. Harry said that the mean is situated between 60 000 and 180 000 which meant that the mean with the lowest value (42 000) has a difference of 30 000 plus. Whereas the mean with the highest value (180 000) has almost a difference between 90 000 to 100 000.

He said if the median is chosen, then the difference is much more compared to the mean. In which he explained that the difference between the median and 180 000 is much bigger than the mean with 180 000. Therefore, Harry felt that the mean is the best to represent the actual salary. He said that to choose the measure that had lesser difference in which he found that the mean had lesser difference as compared to the median. However, Harry looked rather unsure with his answer and complained that he felt it was hard for him to explain. The above Excerpt H9 describes Harry's points on this.

Summary.

Table 4.71

Subject matter knowledge in handling bias

Task	Statistical element	Conclusion
Task 8	Identifies and summarises data using mean	Harry identified and calculated the mean by adding all the values in the data and then divided the total with the total number of data.
	Identifies and summarises data using median	Harry identified the median by first arranging the data in ascending order. He later found the middle data in the arrangement as the median.
	Identifies and summarises data using mode	Harry identified the mode by first finding the frequencies for each data point and then the data point with the highest frequency was identified as the mode.
	Best data representation measure	Harry chose the median as the best data representation measure. He based this by comparing the median with the smallest and the biggest data in the arrangement after eliminating 15.3. He found the difference between the median and these values are quite balanced. Harry said that the mean is affected by the extreme data and is different from the rest of the data which is why the mean is not suitable in representing the data.
Task 9	Identifies and summarises data using mean	Harry identified and calculated the mean by adding all the values in the data and then divided the total with the number of data.
	Identifies and summarises data using median	Harry identified the median by first arranging the data in ascending order. He later found the middle data in the arrangement as the median.
	Identifies and summarises data using mode	Harry mentioned that there is no mode in this situation because every data had only one frequency. Therefore, there was no data which carried the highest frequency.
	Best data representation measure	Harry said that in this situation, he would choose the mean to represent the data. He explained that the median carried a bigger difference with the smallest and the biggest value in the data. He added that the mean had a smaller difference as compared to the median. However, unlike the earlier task here Harry did not exclude the extreme data first before looking at the difference between the measure and the data.

Subject matter knowledge of measures of central tendency in problem solving.

Task 10. Initially Harry calculated the overall GPA by adding the GPAs for both the colleges and then divided the total with two. However, later he realised that the given GPAs are the respective average for all the involved semesters in those colleges.

Table 4.72

Excerpt H10

R: How did you get the overall GPA?
S: GPA for college A 3.2 plus GPA for college B 3.8 divided by 2 because is two different colleges. So, the answer is 3.5.
R: Can you explain further?
S: [Thinks for a while] [Realizes something is wrong, thinks further and tries to work out further] To determine the overall GPA, because 3.2 is based on two semesters and 3.8 is based on three semesters. So, to find the overall GPA, I will use 3.2 times 2 because of the two semesters plus 3.8 times 3. And then divide with the total number of semesters between the two colleges, so two plus three is five. So, 17.8 divided by 5. So, the overall GPA is 3.56.
R: You gave two different answers, which is it?
S: It is 3.56. The earlier one, I did not see the number of semesters involved.
R: Why did you times 3.2 by 2?
S: Because 3.2 is based on two semesters and 3.8 is based on three semesters. That means this value (pointed at 3.8) is the total divided by three. And this value (pointed at 3.2) is obtained after the total is divided by two semesters. We must get the total and then divide with the total semesters. So, the GPA value we get is 3.56.
R: Can you explain further?
S: We find first the total for both colleges for all the semesters and then the total for both colleges are added up and then we divide with the total number of semesters. We want to find the overall GPA for all semesters regardless the college.

Initially, Harry mentioned that the overall GPA is 3.5 which was calculated by adding the GPAs of both colleges and then the total was divided with 2 because of two different colleges. However, after Harry was probed to explain further, he showed some uncertainty. He took some time to think and realized a mistake in the GPA calculation earlier. After which he explained that the GPA for college A which is 3.2 is based on two semesters and GPA for college B which is 3.8 is based on three semesters.

Therefore, Harry now claimed that the overall GPA is found by adding the total grade point for each college which is calculated by multiplying the GPA for each college with the number of semesters involved. Harry elaborated that he would times 3.2 with 2 because of two semesters for college A and 3.8 with 3 because three semesters for college B. Then he would add the total grade point calculated for both colleges and divided the overall total with the total number of semesters which is five. He mentioned that it would be 17.8 divided by 5 and the overall GPA is 3.56.

Realizing that Harry had actually given two different answers, the researcher asked Harry which one his answer is. Harry mentioned that it is 3.56 because earlier he did not realize that for each college there was also number of semesters involved. Harry elaborated that the GPA 3.2 is multiplied with 2 because 3.2 is based on two semesters and the same for GPA 3.8 is based on three semesters. He mentioned that 3.2 is actually obtained from the total grade point for college A divided with two semesters and 3.8 is from the total grade point for college B divided with three semesters. Thus, he said that to get the overall GPA 3.56, the overall total grade point is divided with the overall total semesters. The above Excerpt H10 illustrates Harry's points on this.

Task 11. Harry calculated the weighted mean when he was asked to find the average TV viewing time for all 100 students. He utilised the backward mean calculation to find the total TV viewing time for each of the groups. The totals were later used in the calculation of the weighted mean.

Table 4.73

Excerpt H11

R: How did you get the average?
S: In order to find the average viewing time for students from both rural and urban, I will times 8 with 25 then 4 with 75. Then I plus both the values (pointed at the products of 8 with 25 and 4 with 75). And then I divided this total with the overall total which is 100. I get the average as 5 hours per weekend.
R: Why did you times 25 with 8?

S: Because the average 8 is for 25 rural students and I want to find the total viewing time for rural students.
R: How about 4 times 75?
S: The same, average 4 is for 75 urban students, so I want to find the total viewing time for urban students.
R: Why you add the total?
S: I want to find the overall for 100 students, so I add the total viewing time for the rural and urban and then I divide with 100 to give me the average viewing time for 100 students.

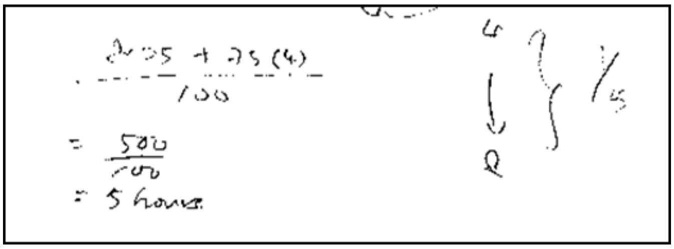


Figure 4.14. Entry H2

Harry mentioned that in order to find the average viewing time for students both rural and urban, he would multiply first 8 with 25 and 4 with 75. Then he said that he would add both the products in which he showed by pointing at his working in the task sheet as shown in Entry H2. The total that he obtained after adding the products is divided with the overall total number of students which is 100 to get the average as 5 hours of TV per weekend.

Harry explained that 25 is multiplied with 8 because he wanted to find the total viewing time for the rural students. Similarly, he said that he multiplied 4 with 75 because he wanted to find the total viewing time for urban students. He mentioned that he added both the products because he wanted to find the average viewing time for the overall 100 students. The total was later divided with 100 to give him the average TV viewing time for 100 students. The above Excerpt H11 illustrates this.

Summary.

Table 4.74

Subject matter knowledge in problem solving

Task	Statistical element	Conclusion
Task 10	Backward mean calculation	Harry solved for the weighted mean without the task explicitly stating to do so. He carried out the "backward" mean calculation, where he obtained the total for each group from the given group means.
	Representative nature of the mean	Harry used the knowledge related to the representative nature of the mean in making sense of why data sets need to be combined in the weighted mean.
	Forward mean calculation	Harry also carried out the "forward" mean calculation to calculate the weighted mean.
Task 11	Backward mean calculation	Harry solved for the weighted mean without the task explicitly stating to do so. He carried out the "backward" mean calculation, where he obtained the total for each group from the given group means.
	Representative nature of the mean	Harry used the knowledge related to the representative nature of the mean in making sense of why data sets need to be combined in the weighted mean.
	Forward mean calculation	Harry also carried out the "forward" mean calculation to calculate the weighted mean.

Subject matter knowledge of measures of central tendency in making inference.

Task 12. Harry said that he would recommend the player to be selected for the All Star team based on their total scores or average. However, Harry also mentioned that game comparison can also be done to see the most eligible player to be selected.

Table 4.75

Excerpt H12

R: If Coach Ahmad can only select one of the two players, who would you recommend he select?
S: I will recommend to select Beng
R: Why Beng?
S: We add all the scores without considering any statistics. For Ramu is 118 and Beng is 131. So Beng has higher total scores compared to Ramu. That means Beng scored higher in the games. But since both of them played six games, the average, the mean score that we get for Beng is 21.83 but Ramu only 19.67. So, from here we can see that Beng should be selected to play in the All Star Team, his chances for scoring for the team is better than Ramu. This is based on the average.
R: Is there any other way that you can use to decide which player to be selected?
S: If we do not want to decide using the total scores and then the average. We can compare each game. Beng did better than Ramu in the first game by 3 scores. In the second game, also Beng did better by 2 scores. Third and fourth game Ramu did better by 2 points and 3 points respectively. Whereas the fifth and sixth game, Beng did better by 2 and 11 points respectively. So, from the

comparison between these two (Beng is given + score and Ramu is given a negative score), we can see 3 and -3 which can be cancelled. 2 and -2 can be cancelled. So, total up (the score difference of each game between the two players) can see that Beng scored more marks compared to Ramu in the difference scores. So, the chances for Beng to score is better than Ramu. I will use this method if asked to use any other method besides the above to select the player. But this is not so accurate.
R: Anything else?
S: No. Only this way is the best, add up the total scores then divide by six.

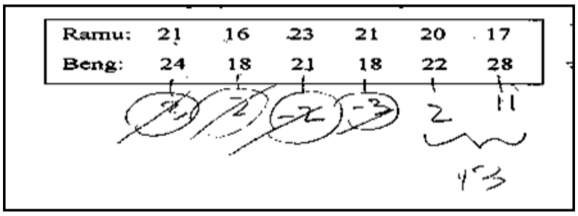


Figure 4.15. Entry H3

Harry mentioned that he would recommend Beng to be selected for the team. He said that he based his decision by adding all the scores without considering any statistics. He compared the total scores for Ramu and Beng. Ramu's total scores was 118 meanwhile Beng's total scores was 131. Therefore, he said that Beng had higher total scores than Ramu which means that Beng scored higher in the games.

Harry further elaborated on the mean score. He said that since both players played six games, the mean score for Beng was 21.83 while Ramu was only 19.67. Thus, he said based on the average or the mean score, Beng should be selected to play for the team because his chances for scoring for the team is better than Ramu.

Besides this, Harry mentioned that the player can also be selected based on each game comparison. He elaborated that Beng did better than Ramu in the first game by 3 scores and was given +3. In the second game Beng again did better by 2 scores and is given +2. In the third and fourth game Ramu did better by 2 and 3 points. Ramu was given -2 and -3 respectively. The final two games, again Beng did better and was given +2 and +11. Based on these + and – scores, the scores were added up. The total was +13, which showed that Beng scored more compared to Ramu. This is shown clearly in Entry H3.

Therefore, Harry mentioned that the chances for Beng to score will be better than Ramu. However, Harry said that he would only use this method if he was asked to use other methods besides total scores or average because he felt that this method is not so accurate. Harry also mentioned that based on the mean score which he said as "add up the scores and divide by six" is the best way to decide on the player to be selected. The above Excerpt H12 illustrates Harry's points on this.

Task 13. Initially, in this task Harry said that a game comparison cannot be done because of there is a difference in the number of games played by both players. Instead, he said that the average comparison should be done. However, later he mentioned that maybe game comparison can be done based on the selection of equal number of games from the player who played more games with the other player who played less. He claimed that based on this, the game comparison can be done. Nevertheless, he said that the selection criteria do play a role here which might influence the player that will be selected.

Table 4.76

Excerpt H13

R: Coach Ahmad would like to select another player now from the group below. Who would he select from this group?
S: I will recommend to select Khan.
R: Why?
S: Since the two of them are not playing the same number of games, Khan played less two games compared to Ali. So, we cannot compare like just now the one to one comparison. We need to total up and then divide by the number of games. We need to find is the average. For Ali is 178 divide by 8 because Ali played 8 games. His average is 22.25. But for Khan after we add up and divide by 6 because he played less two games. Khan's average score is 23. So, which means that Khan is better than Ali. So, I will recommend to select Khan.
R: Are there any other ways that you can use to select the player?
S: Based on the scores, choose the higher scores for six games for Ali. But I am not sure if this method is correct. But if we use this, maybe Khan might not be chosen because when we do game to game comparison on each set for six sets. Then Ali might be selected. Khan got 16 but for Ali among the games selected based on the higher scores, there is no 16. I mean Coach can choose best six among Ali and for Khan there is only six games so maybe Ali might stand a better chance to be chosen. However, if the Coach considers different set selection then Ali might not stand the chance. But it all depends on the condition used to do the selection.

R: Compared to the method just now and the average, which method is the best to select the player?
S: As for me, I will use the average to decide the player and by using the division by eight for Ali and division by six for Khan.
R: Why?
S: Because is fair enough. I let you play and score and then consider all the games that you played. So is fair for Ali and Khan. Each game plays a role in the selection.

Harry mentioned that he would recommend Khan to be selected. He mentioned that unlike the earlier task which had equal number of games for both players, in this task the number of games played by both players are not the same. Therefore, he said that the one to one game comparison cannot be done. Instead, he said that "to total up and then divide by the number of games" or calculate the average. For Ali, he mentioned that the average is 22.25 calculated from the division of 178 with 8 because Ali played eight games. Whereas Khan's average is 23 obtained from the division of his total with 6 because he played only six games, less two games from Ali. Since Khan's average was better than Ali's, therefore Harry said that he would recommend Khan to be selected.

Even though Harry mentioned earlier that for this situation a one to one game comparison cannot be done but later Harry said that it can be done if the scores are selected equally. He said that six top scores for Ali can be chosen for the game comparison. However, he was unsure if this method would be accurate. He explained that if six top scores are selected for Ali and compared with the six scores that Khan already has, then maybe Khan might not be selected. Instead Ali might be selected.

If Ali's scores are based on the top six than his lowest score is higher than Khan's lowest score 16. He explained that Ali's chances of being selected or not depends on the coach's selection criteria. He said if the coach had used a different set for the selection than maybe Ali might not stand a chance. Therefore, Harry added that it all depends on the condition or the selection criteria used.

Nevertheless, when Harry was asked on the best method to select the player, he mentioned that the average is the best with the division considered in the average is over eight for Ali and over six for Khan. Harry explained that the average is considered as fair because all the games that the player had played is taken into account. He said considering the average is also fair for Ali and Khan. Harry also said that in the average each game played by the player had a role in the selection. The above Excerpt H13 describes Harry's points on this.

Task 14. Harry had three different ways to look at the class performance. He compared the graphs, the mode and the total number of corrects to decide on the class that performed better.

Table 4.77

Excerpt H14

R: Based on the graphs below, which class did better?
S: 5A did better.
R: Why?
S: Because from the graph for 5A, you can see that the number of correct 6 got 3 students and then 5 also got 3 students and for the worst that is 3 corrects there is only 1 student. But for 5B, 6 corrects there is only 1, 5 corrects is 2 and the worst which is 3 corrects there is 2. This means when we draw a line, you can see that 5A positively skewed. The more students who got more number of corrects is better. But this one (points at 5B) is slightly negatively skewed.
R: Is there any other way that you can use to choose the better class?
S: We can use the mode to find which class did better. For this one (points at 5B) the mode is 4 but this one the mode is 5 or 6. So obviously 5 or 6 corrects is much better than 4.
R: Are there any other ways to choose the class that did better?
S: Mmm… I will try to multiply.
R: Multiply what?
S: The number of corrects multiply with the number of students. For example, 6 times 3 is 18 so there is 18 correct answers here. And then 5 times 3 is 15 and so. Then we try to total up to get the overall total number of corrects. Of course, the class that did better, the overall total number of corrects will be higher than the other class. After we multiply the number of corrects with the number of students and we total up to get the overall total number of corrects, for 5A gets 44 and for 5B is 40. So 5A did better.
R: Anything else?
S: No either graph or individually.

Harry said that 5A performed better. He explained that from the graphs, 5A had more students who scored more number of corrects as compared to 5B which had lesser students who scored more number of corrects. He stated this in his example, for 5A, 3 students obtained 6 number of corrects, another 3 students got 5 number of corrects and only 1 student got 3 number of corrects. On the other hand, 5B had only 1 student who obtained 6 number of corrects, 2 students got 5 number of corrects and 2 students got 3 number of corrects. He said that if a line is drawn on each of the graphs, 5A is positively skewed which means more students scored better. On the other hand, he said that 5B is slightly negatively skewed. Thus, 5A did better than 5B.

Apart from this, Harry mentioned that the mode can be used to determine the class that performed better. He pointed that the mode for 5B is 4 whereas the mode for 5A is 5 or 6. So he said that obviously 5 or 6 is better than 4, therefore 5A did better. When he was asked further on any other ways to determine the class that performed better, Harry mentioned through multiplication.

He elaborated that to multiply the number of corrects with the respective number of students. For instance, 6 number of corrects is multiplied with 3 students to get 18 number of corrects and so for the rest. He then said that all the products of the multiplication would be added to get the overall total number of corrects for both the classes. Harry said that the class that did better would have a higher overall total number of corrects than the other class. He concluded that 5A did better because the overall total for 5A was 44 which is higher compared to 5B which was only 40. The above Excerpt H14 describes Harry's points on this.

Task 15. Harry mentioned that when comparing two classes of different number of students, the performance of the classes can be determined by looking the skewness of the graphs.

Table 4.78

Excerpt H15

R: Based on the graphs below, which class did better?
S: If compare these two graphs, firstly we can clearly see that the number of students are different, one (5C) is 36 and the other one (5D) is 21. Since the number of students are different, the best method to choose is straight away draw the graph. (Draws the curves on the graph). So just like the previous question, I will choose 5D.
R: Why?

S: I choose 5D even though the number of students is much lesser than this one but this one (points at 5C). We can see that the graph is skewed to the right which means that the number of students who got more number of corrects for example 6, 7, 8, 9 is more compared to 5C. We ignore the total number of students and look at the graphs only.
R: Besides this, is there any other ways to determine which class did better?
S: Here we cannot base on the total number of questions correct because there is a difference in the number of students. 5C has more students than 5D, so totally out.
R: How about measures of central tendency?
S: According to my understanding, I will not decide based on the measures of central tendency but I will use the graphs to decide. Look at the skewness to decide. I will not use mean, mode or median because the difference in the number of students. If the number of students is the same, then we can consider the mode. Since the number is different so I will not use the mean, the mode or the median to decide on which class did better. I will decide based on the graphs.

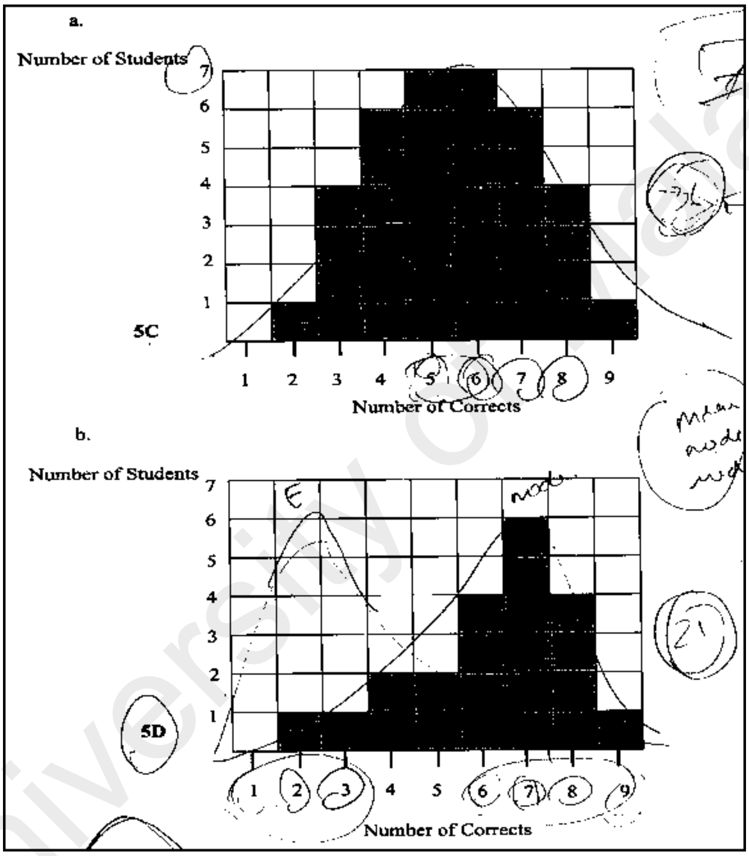


Figure 4.16. Entry H4

For this task, Harry claimed that 5D did better. He compared the graphs of both the classes and said that the first thing which was obvious to him was the number of students for both the classes were different. He said that 5C had 36 students whereas 5D had only 21. Since the number of students were different, so he claimed that the best method to determine the class that performed better is by drawing the curves on the graphs. This shown in Entry H4. Based on this, he said he would choose 5D.

He explained that he chose 5D as the class that performed better even though the number of students in 5D was much lesser than the number of students in 5C. This was due to the graph for 5D was skewed to the right which he said indicated that there were more students who scored more number of corrects such as 6, 7, 8, and 9 as compared to 5C. He said that he completely ignored the total number of students and just focused on the graphs. He elaborated that for this situation, the comparison on the total number of questions correct cannot be used because of the difference in the number of students.

Harry mentioned that he would not also use the measures of central tendency to decide on the class that performed better. He said that he would use the graphs to decide by looking at the skewness. He explained that this was because the difference in the number of students in both the classes. He added that if the number of students happen to be the same, then he would consider. But since the numbers are different, therefore he said that he would not use the mean, the mode or the median to decide the class that performed better. The above Excerpt H15 describes this.

Summary.

Table 4.79

Subject matter knowledge in making inference

Task	Statistical element	Conclusion
Task 12	Summarises equal sized numerical data using measures of central tendency	Harry summarised the equal sized numerical data using the mean.
	Utilises the appropriate measure to compare and draw conclusion about two equal sized numerical data	Harry compared the two-equal sized numerical data using the total scores and the average. He said that both can be used to decide on the player to be selected. He also said that a one to one game comparison can also be done to select the player. However, among all these ways that he mentioned he said that the one to one game comparison is not so accurate and he would only use it if he was asked to use a method apart from the total scores and the average. According to Harry, the best method to choose the player would still be using the average.

Task 13	Summarises unequal sized numerical data using measures of central tendency Utilises the appropriate measure to compare and draw conclusion about two unequal sized numerical data	Harry summarised the unequal sized numerical data using the mean. By adding all the game scores and then divided the total with the number of games each player played. Harry compared the average score for both the players and decided on the player to be selected. He also mentioned another way to select the player would be a one to one game comparison which initially he said cannot be done. However, he said that this can be done if it is based on equal game sets. He said that for the player who played more games, six games would be selected to be compared with the six games of the other player. However, Harry said that the selection criteria do play a role here because in a way it also determines the player that likely to be chosen. Nevertheless, Harry mentioned that the average is still the best method to be used because he claimed that it gives a fair selection.
Task 14	Summarises equal sized graphical data using measures of central tendency Utilises the appropriate measure to compare and draw conclusion about equal sized graphical data	Harry summarised two equal sized graphical data using the mode and the total scores. Harry compared the two-equal sized graphical data using the skewness of the graphs. He said that the graph which is positively skewed is better. He also used the mode and said that the class that had bigger mode performed better. Harry also got the total number of corrects for both the classes in which he multiplied the number of students with the number of questions correct. He compared both the totals and decided on the class that performed better.
Task 15	Summarises unequal sized graphical data using measures of central tendency Utilises the appropriate measure to compare and draw conclusion about unequal sized graphical data	Harry did not summarise the two-unequal sized graphical data using any of the measures of central tendency. Harry did not utilise any of the measures of central tendency to compare and draw conclusions about two unequal sized data. In fact, he said that since the two data are of unequal size, therefore the measures of central tendency cannot be used. Instead, he used the skewness of the graphs to decide on the class that performed better.

Level of subject of measures of central tendency.

Level of subject matter knowledge of measures of central tendency with reference to context. In Task 1, Harry mentioned that the average was found by adding up all the children from all the families and then divided the total with the number of families. He also showed the mean calculation in his example. Harry knew that for the context of the given data which was referring to the number of children, the ideal average was not logic. However, he said that any division involving two non-divisible numbers can result in a decimal average. Therefore, Harry had presented the knowledge of the mean as the average (P) which is considered as appropriate (A).

In Task 2, Harry had utilised the idea of mode without the task explicitly stating so. He found the frequency for each shoe size and had identified the mode as 5 based on the highest demand. Therefore, Harry's idea of the mode was found present (P) and correct (A).

Harry said that the mode can be a form of data representation. He explained that a lot people wore size 5 but there were also other shoe sizes to be considered too. Therefore, Harry's knowledge of the mode as a form of data representation was found to be present (P) and correct (A).

Harry mentioned that he would use the mode as the average because it indicated the shoe size worn by most customers. According to Harry, the average here is basically to place the order for the shoes and in this situation the mode can be used as the average shoe size. Therefore, Harry's knowledge of the mode as the average was found to be present (P) and appropriate given to his justification (A).

In Task 3, Harry identified the mode as student D. He said that he based this on the highest frequency from the data. When the task was modified where the number of comments for student D was changed from 22 to 6, Harry still identified the mode as student D which he said that he based on the highest frequency. Harry had identified the mode incorrectly. He had taken the highest number of comments as the highest frequency. He identified the mode based on the category that carried the highest frequency. Therefore, Harry's idea of the mode was marked to be not present (NP) and the lack is considered problematic (I).

Harry mentioned that he would not utilise the mode to represent the number of comments made by the students. He said that some students hardly gave any comments and some gave many comments, so the mode cannot be taken to represent the comments given by all the students. He added that the mode was based on the comments of one student. Instead, he said that the mean is more suitable because the mean considers all the comments given by all the students in the class. Harry lacked the knowledge that the mode can be utilised as a form of data representation. Therefore, Harry's knowledge of the mode as a form of data representation was found to be not present (NP) and the lack was considered problematic (I).

In Task 4, Harry mentioned that the average means that most of the students watched 3 hours of TV per day. However, his later explanations revealed that his knowledge of the average is the mean when he elaborated on the mean calculation in his example. Harry said that in the calculation of the average, all the data including the extreme data need to be considered. He said that the information on the extreme data can deviate the average and the average might not be accurate.

Harry also mentioned that the average can be based on the mode if the sample is big enough. However, he said that the average cannot be based on the median because the median referred to the middle value in the 24 hours' time system. So according to him, the median was 12 hours and not 3 hours.

Harry was not aware that the average based on any data involving human population such as the Malaysian primary school students is normally based on the median because the median is definitely a more robust measure when dealing with data that contained extreme data. Therefore, Harry did not present the knowledge of the median as the average (NP). However, Harry's knowledge of the mode can be considered as the average when the sample size was big enough is considered appropriate (A).

In Task 5, Harry found the median by rearranging all the weights from the smallest to the biggest number. He found the median weight for nine children as 19 which was the middle weight in the rearrangement. He also found the median weight for 10 children after 43 kg was included in the data. He got 19 as the median after he added the two middle numbers and then divided the total by two. Harry's knowledge of the idea of the median was found to be present (P) and accurate (A).

Harry said that the mean is not adequate to represent the weight of all ten children. He explained that this was because the weight of the tenth child which was much heavier compared to the rest. Instead Harry said that the median is more suitable because it represented the data. However, Harry's idea of the suitable representation measure was based on the difference between the median with the smallest and the biggest values in the data. Harry excluded the extreme data when he did the median comparison. Whereas, Harry did not exclude the extreme data from the mean calculation. Therefore, he claimed that there were bigger differences when the mean was compared with the smallest and the biggest values in the data which did not exclude the extreme data.

Harry's knowledge related to the robustness of the measures of central tendency was found to be present (P). He knew that the median is a more robust measure in the presence on an extreme data. However, his justification towards the median being more robust was found to be in appropriate (I).

In Task 6 and Task 7, Harry constructed data sets based on the given average value and the conditions imposed. In Task 6, Harry placed all the prices of the seven bags of crisps as 27. He said that this is because the average is 27 when multiplied with 7 and divided with 7 again will give him 27 again. Since he was allowed to use 27, therefore he placed 27 for all seven bags of crisps. Therefore, Harry's knowledge of constructing a data set based on the idea of average was present (P) and found to be appropriate (A).

In Task 7, Harry calculated the total price as 189 by utilising the backward mean calculation. He said that regardless whatever price that he placed, he made sure that the total of these prices was 189. He said that he randomly placed the values. He chose 20 and 29 as the first two values. He subtracted these values from the total and the balance he divided with 5 in which gave him 28. He utilised this value for the rest of the prices. He said any values is possible. In fact, it can also be as low as 1. Thus, Harry's knowledge of constructing a data set based on the idea of average for Task 7 was found to be present (P) and appropriate (A).

Table 4.80

Coding rubrics for determining level of Harry's subject matter knowledge of measures of central tendency with reference to context

Task	Subject Matter Knowledge of Measures of Central Tendency with reference to Context																															
	Mean as average				Mode as average				Median as average				Idea of mode				Idea of median				Mode as data representation				Robustness of measures				Idea of average			
	P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP	
	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I
Task 1	X																															
Task 2					X								X								X											
Task 3																								X								
Task 4											X																					
Task 5																	X									X						
Task 6																													X			
Task 7																													X			
	1				**1**						**1**		**1**				**1**				**1**								**2**			

Legend: P = Present NP = Not Present A = Appropriate I = Inappropriate

Harry's Percentage of Subject Matter Knowledge of Measures of Central Tendency with Reference to Context

$$= \frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)} \text{ X } 100\% = \frac{\mathbf{8}}{\mathbf{11}} \times \mathbf{100\%} = 72.73\% \text{ (Medium)}$$

Level of subject matter knowledge of measures of central tendency in handling bias. In Task 8, Harry calculated the mean for all nine data. He added up all the nine data and then divided the total with nine. He got the mean as 7.18. Therefore, Harry not only identified that the mean can be obtained but also summarised the data using the mean by calculating the mean value. Therefore, Harry's knowledge of identifying and summarising the data using the mean was found to be present (P) and correct (A).

Harry said that the median is 6.15 which was the middle or the fifth number after the data was arranged in the ascending order. Thus, Harry's knowledge of identifying and summarising the data using the median was found to be present (P). However, Harry had arranged the numbers incorrectly where he placed 6.15 after 6.2 in the ascending arrangement. Thus, although he identified the median accurately as the fifth data but the median value was incorrect (I).

Harry identified the mode as 6.3. According to him 6.3 occurred thrice. Therefore, Harry's knowledge of identifying and summarising the data using the mode was found to be present (P) and correct (A).

Harry mentioned that the median is the best measure to represent the actual weight of the object. He explained that after excluding 15.3, the median 6.15 is considered as the most balanced value because it is in the middle of the smallest value which is 6.0 and the biggest value which is 6.3. He said that the mode cannot be considered because the biggest value in the data is the mode itself. Whereas according to him, the mean which is 7.18 is already affected by the extreme data and does not represent the rest of the values in the data which are around 6. Harry's knowledge related to the best representation measure is found to be present (P) because the median can be used as the best representation measure. However, his justification towards the suitability of the median as the best representation measure was found to be inappropriate (I).

Similarly, in Task 9, Harry calculated the mean for all five data. He added up all the five data and then divided the total with five. He got the mean as 76 800. Harry not only identified that the mean can be obtained but also summarised the data using the mean by calculating the mean value. Therefore, Harry's knowledge of identifying and summarising the data using the mean was found to be present (P) and correct (A).

Harry found the median as 54 000. He arranged the data in ascending order first and then identified the median as the middle number in the arrangement. Thus, Harry's knowledge of identifying and summarising the data using the median was found to be present (P) and correct (A).

Harry mentioned that the data in this task does not contain any mode because each data exists as a value by itself. Therefore, Harry's knowledge of identifying and summarising the data using the mode was found to be present (P) and correct (A).

Harry mentioned that the mean is the best measure to represent the actual salary of the company. According to him, the mean 76 800 had a lesser difference when compared to the smallest and the biggest values in the data, unlike the median which carried a much bigger difference. Thus, Harry's knowledge related to the best representation measure is found to be present (P) because the mean can be used as the best representation measure. However, his justification regarding the best representation measure was found to be inappropriate (I).

Table 4.81

Coding rubrics for determining level of Harry's subject matter knowledge of measures of central tendency in handling bias

Task	Subject Matter Knowledge of Measures of Central Tendency in Handling Bias															
	Identifies and summarises data using mean				Identifies and summarises data using median				Identifies and summarises data using mode				Best data representation measure			
	P		NP		P		NP		P		NP		P		NP	
	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I
8	X					X			X					X		
9	X				X				X					X		
	2				**1**				**2**							

Legend: P = Present NP = Not Present A = Appropriate I = Inappropriate

Harry's Percentage of Subject Matter Knowledge of Measures of Central Tendency in Handling Bias

$$= \frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)} \text{ X } 100\% = \frac{\mathbf{5}}{\mathbf{8}} \times \mathbf{100}\% = 62.5\% \text{ (Medium)}$$

Level of subject matter knowledge of measures of central tendency in problem solving. In Task 10, Harry carried out the backward mean calculation to obtain the total grade point average for College A based on two semesters as 3.2 times 2. He also found the total grade point average for College B based on three semesters as 3.8 times 3. Therefore, Harry's knowledge of performing the backward mean calculation to obtain the total from the mean is found to present (P) and correct (A).

Harry explained that the mean which was given for each college was for the number of semesters involved in that particular college. He also added the total grade points of both colleges and the total semesters of both colleges before he proceeded in the weighted mean calculation. Thus, Harry's knowledge related to the representative nature of the mean is found to be present (P) and appropriate (A).

Finally, Harry carried out the forward mean calculation to find the weighted mean. He divided the total grade points of both colleges with the total semesters of both colleges and found the weighted mean as 3.56. Therefore, Harry's knowledge of the forward mean calculation is found to present (P) and appropriate (A) despite the task not explicitly stating to find for the weighted mean.

In Task 11, Harry performed the backward mean calculation and obtained the total hours of TV watched per weekend for both rural and urban groups. He said that he multiplied the number of students in each group with the group average. Harry found the total hours of TV watched for rural group by multiplying 25 with 8. Similarly, for the urban group he multiplied 75 with 4. Thus, Harry's knowledge of performing backward mean calculation was found to be present (P) and appropriate (A).

Harry explained that he multiplied the average of each group with the number of students in each group because the given average was for the group. He also mentioned that the totals of both the groups were added and the total number of students of both groups were found before he proceeded with further calculation because he wanted to find the overall average for all the students. Thus, Harry's knowledge related to the representative nature of the mean is found to be present (P) and appropriate (A).

Finally, Harry carried out the forward mean calculation to find the weighted mean. He divided the total number of hours of both groups with the total number of students and found the average as 5. Therefore, Harry's knowledge of the forward mean calculation is found to present (P) and appropriate (A) despite the task not explicitly stating to find for the weighted mean.

Table 4.82

Coding rubrics for determining level of Harry's subject matter knowledge of measures of central tendency in problem solving

Task	Subject Matter Knowledge of Measures of Central Tendency in Problem Solving											
	Backward mean calculation				Representative nature of the mean				Forward mean calculation			
	P		NP		P		NP		P		NP	
	A	I	A	I	A	I	A	I	A	I	A	I
10	X				X				X			
11	X				X				X			
	2				**2**				**2**			

Legend: P = Present NP = Not Present A = Appropriate I = Inappropriate

Harry's Percentage of Subject Matter Knowledge of Measures of Central Tendency in Problem Solving

$= \frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)} \text{ X } 100\% = \frac{\mathbf{6}}{\mathbf{6}} \times \mathbf{100}\% = 100\% \text{ (High)}$

Level of subject matter knowledge of measures of central tendency in making inference. In Task 12, Harry calculated the mean scores for both the players by adding all the scores and divided the total with the six set of games that they played. He found that the mean score for Ramu was 19.67 and the mean score for Beng was 21.83. Therefore, Harry's knowledge of summarising equal sized numerical data using the measures of central tendency in this case the mean without the task explicitly stating to do so was found to be present (P) and appropriate (A).

Harry compared the mean scores of both the players and decided on the player that he would recommend to be selected for the team. Since Beng's average was higher than Ramu's, he said that Beng would be recommended. Harry also mentioned that the player can also be selected by comparing the total scores or each game set. The player that had higher total or the one that had won the most sets would be selected. However, Harry mentioned that the average comparison is the best way to compare and decide on the player to be selected. Thus, Harry's knowledge of utilising the appropriate measure of central tendency to compare and draw conclusion on two equal sized numerical data was found to be present (P) and appropriate (A)

In Task 13, Harry summarised the data using the mean or the average. He calculated the average by adding all the scores for each player and then divided the total scores with the number of game sets that they played. He said for Ali, his total scores were divided with eight because he played eight game sets. Whereas for Khan, his total scores were divided with six because he played six game sets. He found that Ali's average was 22.25 while Khan's average was 23. Therefore, Harry's knowledge of summarising unequal sized numerical data using the measures of central tendency in this case the mean without the task explicitly stating to do so was found to be present (P) and correct (A).

Harry compared the average scores of both the players and decided to recommend Khan because his average score was higher than Ali's. He said that the average was a fair measure because all the game sets that the player played were taken count. He added also that each game played a role in the selection. Although Harry mentioned that the game comparison based on the selection of six games for the player that played more sets can be done but this method is considered not so accurate. According to him, the selection criteria has an influence on the player to be selected. So, Harry said that the average was still the best method to select the player.

Thus, Harry's knowledge of utilising the appropriate measure of central tendency in this case the mean to compare and draw conclusion on two unequal sized numerical data was found to be present (P) because the mean is the most suitable measure to be used to compare and draw conclusions based on unequal sized data. Harry's knowledge of utilising the mean here was considered to be appropriate too (A).

In Task 14, Harry summarised the data using the mode. He found that the mode for 5B is 4 and the mode for 5A is 5 or 6. He also found the total number of corrects for both classes by multiplying the number of students with the number of questions correct. Since 5A had a total of 44 as compared to 5B which had a total of 40, thus he said that 5A performed better. Harry had showed that he utilised the mode to summarise the equal sized graphical data. Therefore, Harry's knowledge of summarising equal sized graphical data using the measures of central tendency in this case the mode without the task explicitly stating to do so was found to be present (P). He had identified the mode correctly too (A).

Harry compared the mode for both the classes and since 5A's mode was higher than 5B, he said that 5A performed better. Harry also compared the graphs and looked at the skewness of the graphs. Although Harry did not calculate the mean of both the classes but he did find the total number of corrects of both classes. Since 5A had a higher total than 5B, he said that 5A performed better. Thus, Harry's knowledge of utilising the appropriate measure in this case the mean to compare and draw conclusion on two equal sized graphical data was not present (NP). He did not use the mean when the mean is the most suitable measure to compare and draw conclusions related to the performance of the classes based on two equal sized graphical data. This absence is regarded as problematic (I).

In Task 15, Harry did not summarise the two-unequal sized graphical data using any of the three measures of central tendency. In fact, he claimed that since the data are of unequal size, therefore the measures of central tendency is not appropriate to be used in this situation. Thus, Harry's did not show the knowledge of summarising two unequal sized graphical data using the measures of central tendency without the task explicitly stating to do so (NP) and this was found to be problematic (I).

Harry compared the skewness of the graphs to decide on the class that performed better. He said that since the number of students in both the classes are different, neither the total number of questions correct nor the measures of central tendency can be utilised to compare and draw conclusions on the performance of the class. Thus, Harry did not show the knowledge of utilising the appropriate measure of central tendency to compare and draw conclusions on two unequal sized graphical data (NP) and the absence is found to be problematic (I).

Table 4.83

Coding rubrics for determining level of Harry's subject matter knowledge of measures of central tendency in making inference

Task	Subject Matter Knowledge of Measures of Central Tendency in Making Inference																															
	Summarises two equal sized numerical data using measures of central tendency				Utilises the appropriate measure to compare and draw conclusion about two equal sized numerical data				Summarises two unequal sized numerical data using measures of central tendency				Utilises the appropriate measure to compare and draw conclusion about two unequal sized numerical data				Summarises two equal sized graphical data using measures of central tendency				Utilises the appropriate measure to compare and draw conclusion about two equal sized graphical data				Summarises two unequal sized graphical data using measures of central tendency				Utilises the appropriate measure to compare and draw conclusion about two unequal sized graphical data			
	P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP	
	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I
Task 12	X				X																											
Task 13									X				X																			
Task 14																	X							X								
Task 15																												X				X
	1				**1**				**1**				**1**				**1**															

Legend: P = Present NP = Not Present A = Appropriate I = Inappropriate

Harry's Percentage of Subject Matter Knowledge of Measures of Central Tendency in Making Inference

$= \frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)}$ X 100% = $\frac{\mathbf{5}}{\mathbf{8}} \times \mathbf{100\%}$ = 62.5% (Medium)

Harry's mean percentage of subject matter knowledge for all four constructs $\frac{\mathbf{72.73+62.5+100+62.5}}{\mathbf{4}}$ = $\mathbf{74.43}$%. Therefore, Harry's level of subject matter knowledge of measures of central tendency is medium.

Case Study Four: Amy

Amy was 24 years 2 months old when she was interviewed. At the point of data collection, she was pursuing a 4-year Bachelor of Science with Education (B.Sc. Ed.) program at a public university. She majored and minored in biology and mathematics respectively. Amy obtained an A in Mathematics and a B+ in Additional Mathematics in her SPM examination (equivalent to O level examination). She also scored a B+ in Mathematics T in her STPM examination.

At the time of data collection, Amy was in the final semester of fourth year studies. She attained 3.46 in the Cumulative Grade Point Average (CGPA) for her first three years of studies in the public university. She does not have any teaching experience prior to this interview. The following sections discuss the findings of Amy's subject matter knowledge of measures of central tendency in each of the four different constructs, namely with reference to context, in handling bias, in problem solving, and in making inference that were emerged from the clinical interview.

Types of subject matter knowledge.

Subject matter knowledge of measures of central tendency with reference to context.

Task 1. Amy's knowledge of the mean is average which is seen clearly in her example where she elaborated on the mean calculation. Amy explained that a decimal average is obtained as a result of non-divisible numbers involved in the mean calculation.

Table 4.84

Excerpt A1

R: How do you think the average was obtained here?
S: You collect the data from each family and then you divide by the number of families.
R: Can you show and explain to me using an example?

S: For example, family A has 5 children, family B has 2 children, family C has 3 children and family D has 4 children. So, you just add this up, equals to 16 children and there are four families here so divided by four. So, the average is 4 children.
R: How can the average be 2.3 and not a counting number like 1, 2, 3, 4.
S: Mmm…because the number of children obtained from the data is not divisible by the number of family. For example, if family A has 2 children, B has 3 children. So, you get the average as 2.5 children. So that means 5 cannot be divided by 2 completely.
R: How would you explain about 2.3 children?
S: [Paused for a while] This is what we call as the mean or the average number. It is not only used in human population but it is used in everything. So, this is just a representation of the average number. Meaning if there are 23 children in 10 families, so the average will be 2.3. Average is obtained by dividing the total number of children with the number families. If the numbers are not divisible then you can get a decimal number and not a counting number like 1, 2, 3 and so.
R: What does average relate to in terms of measures of central tendency?
S: Mean
R: Can average refer to any other measures of central tendency?
S: No. It can only refer to mean.

Amy mentioned that the average is obtained from the data collected for each family and then divided with the number of families. She elaborated in her example that if there are four families A, B, C, and D and these families have 5, 2, 3, and 4 children respectively, to add up the number of children for all four families and then divide with the total number of families 4. This will give an average of 4 children.

Amy explained that the average can be a decimal number and not a counting number if the total number of children obtained from the data is not divisible with the total number of families. She further elaborated this fact with an example where if family A has 2 children and family B has 3 children then the average will be 2.5 a decimal number. She mentioned that since the total number of children is 5 which is not divisible completely with the total number of families that is 2.

Amy mentioned that 2.3 is the mean or the average which is not only used for representations related to human population but also for other types of data. It happened to be that here, for an example the number of children involved is 23 for 10 families. Therefore, the decimal average 2.3 is obtained by dividing the total number of children with the total number of families. The average happened to be a decimal number and not a counting number like 1, 2, 3, and so because two non-divisible numbers were divided. Amy claimed that the average is related to the mean and not any other measures of central tendency. The above Excerpt A1 illustrates Amy's points on this.

Task 2. Amy utilised the idea of mode without the task explicitly stating to do so. She mentioned that she would use the mode to represent the shoe sizes if the mode and the mean are around the same value. Otherwise, she said that she would not use the mode because there are other data to be considered too.

Table 4.85

Excerpt A2

R: How would you place the order for the shoe sizes in your shop?
S: I will place the order according to the highest purchase. That most people buy. I will purchase those shoe sizes more.
R: Based on this task, how would you place the order?
S: I would purchase more on size 5 and the other sizes I will purchase lesser. So size 5 and then size 7 and so.
R: How do you base your purchase?
S: I purchase based on the frequency.
R: If you want to choose one female shoe size, which shoe size would you choose?
S: Size 5.
R: Why?
S: Because of the highest frequency which shows that most female can fit this size.
R: Can the chosen shoe size represent the female shoe size in your shop?
S: Actually, initially I was thinking no because there are other data that you need to include also but then after thinking about it I think can.
R: Why?
S: Because the mode and the mean is around the same value.
R: What are the values?
S: Here the mode is obviously 5 and the mean I calculated is 5.5.
R: What if the mode and the mean is not around the same value, would you still use the chosen shoe size to represent the female shoe size?
S: [paused a while and looked unsure] to represent means what?
R: To represent means to [mewakili explains in Bahasa]
S: No. Because the other data plays an important rule too such as size 4, size 6 and so.
.
.
.
R: What would you use as the average shoe size here?
S: 5.5
R: Why?
S: Because the question is asking for the average or the mean.
R: How did you get this value?
S: 5.5 is obtained by calculating. I total all the given shoe sizes which I get 33 and then I divide with 6 for the different shoe sizes and I get 5.5.

Amy mentioned that she would place the order for shoe sizes according to the highest purchase. She said that the shoe sizes that most people buy she would purchase more on those sizes. She elaborated that she would purchase more on size 5 as compared to the other shoe sizes in which she said would be purchased lesser. She added that the purchase will be based on the frequency.

Amy mentioned that if she were to choose one shoe size, it would be size 5 because this size had the highest frequency that indicated most ladies fit this size. However, Amy had doubts on whether to consider the chosen shoe size to represent the female shoe sizes in her shop. She mentioned that initially she wanted to say no but she changed her mind after considering the fact that the mean value is also around the mode value which is 5. She said that the mode is 5 and the mean that she calculated was 5.5.

However, after Amy was probed further with another situation what if the mean and the mode are not around the same value, she claimed now that she would not use the mode to represent the shoe sizes in her shop. She said that there are also other data which is important to be considered.

Finally, Amy mentioned that she would use the mean that she calculated as the average shoe size. She mentioned clearly in her response saying that the question is asking for "the average or the mean". After when Amy also described the mean calculation for how she obtained 5.5. She explained that she totalled up all the shoe sizes in which she got 33. Later she divided the total with 6 for six different shoe sizes. The above Excerpt A2 describes Amy's points on this.

Task 3. Amy identified the mode based on the highest number of comments. Amy believed that if a measure is considered to represent a set of data, therefore the measure has to consider all the data in the set. Thus, she mentioned that the average is a better measure to represent a set of data.

Table 4.86

Excerpt A3

R: What is the mode here? S: Student D with 22 comments R: How did you obtain this value as the mode? S: Because it has the highest number of comments as compared to the other students. R: If you were to represent one value for the number of comments made by the students on that day, would you use this value? S: No R: Why? S: Because you can see that this number is considered very extreme compared to the others. So, if you want to set a number which represent the number of comments, I think it is better if use average. R: How do you find the average? S: Add all up and divide by 8 R: If I were to change this value to 6, what is the mode now? S: Still student D R: How did you obtain the mode? S: Based on the highest number of comments. R: And if you were to represent one value for the number of comments made by students on that day, would you use this value now? S: No R: Why? S: Because to represent the number of comments that means you are going to represent or take account for the number of comments made by all the students in the class. It is better to involve everyone. It is better to use the mean.

Amy identified the mode as student D because student D had the highest number of comments which is 22. Amy mentioned that she would not use the mode to represent the number of comments made by students on that day. She explained that the number of comments for the mode was considered very extreme compared to the other number of comments. Instead she added that the average would be better to represent the number of comments made by students on that day. Amy described that the average can be found by adding all the number of comments and then divide the total with 8 for eight students.

Amy was probed further with the task modified in which the original number of comments for student D was changed from 22 to 6. Amy still identified the mode as student D and she still based the mode on the highest number of comments. However, Amy still felt that the new mode cannot be used to represent the number of comments made by the students on that day. She felt if a measure is taken into account to represent the number of comments made by the students than that particular measure must consider the number of comments made by all the students in the class. She said that it is better to involve everyone. Thus, according to Amy, it is better to use the mean. The above Excerpt A3 describes Amy's points on this.

Task 4. Amy's understanding of the average is the mean. She believed that in finding the mean an outlier can be discarded but for the average all the data including the outlier are taken into account.

Table 4.87

Excerpt A4

R: What does average mean in this sentence? S: Average takes into account all the Malaysian primary school students. R: What do you mean it takes into account all Malaysian primary school students? S: Meaning the average is looking at for all Malaysian primary school students R: How do you think they got this average? S: They have to interview each student and ask them the number of hours, they spend watching TV and then add the number of hours and divide by how many students they interview. R: Show me with an example S: Like student A watch for 2 hours, student B for 5 hours, student C for 3 hours, student D for 8 hours, student E for 3 hours. So, you add up [Writes down the example and calculates] and then divide by 5 because 5 students. R: What does the calculation show? S: Mean calculation R: What about the information on a small group of students who watched more than 12 hours of TV per day. How do they deal with this information? S: They add up together and divide together. R: How? S: Maybe discard [unsure]. I think they would not because here mention that an average of what so. Average should take into account of all data that they have collected. R: Earlier on you mention about discard. What do mean by this? S: When there is an outlier, you can discard. But cannot when calculating the average. R: Can you explain further? S: [Paused and thinks for a while] Not sure. {Seems to be unsure on how to deal with an outlier in the data such as when to discard} When the data is too far away. For example, when you interview most of the students like for example 1 or 2 or 3 hours and then if there is a group of students watch like 10 or 20. So if you know that the information is not right. For example, if a school student tells you that he watches for 18 hours, you know it is not possible because 24 hours and you go to school for at least 8 hours and then if you know that the information is not correct then maybe you can discard. R: How about here? S: Mmm… no. Cannot discard R: Why? . . . S: [Thinks for a while] No. Should not discard because the average is for all Malaysian students. R: Is there any other way they would have found the average? S: No. I know only this. To add all the number of hours and divide by the number of students.

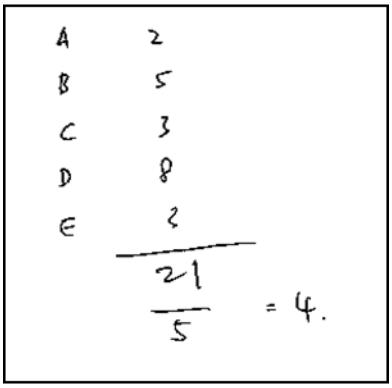


Figure 4.17. Entry A1

Amy said that the average takes into account for all Malaysian primary school students. She elaborated that the average was obtained from the data gathered through interview where the students were asked on how many hours of TV they spend watching. After when the number of hours that these students watched were added and then divided with the total number of students who were interviewed.

This was also seen clearly in Amy's example in Entry A1. She mentioned that if there were five students and each of the students watched 2, 5, 3, 8, and 3 hours of TV respectively. The number of hours were added up and then divided with the number of students. Amy mentioned that the above calculation is referring to the mean calculation. Amy said that the information on a group of students who watched more than 12 hours of TV would also be included in the calculation.

However, after she was probed, Amy mentioned maybe the information was discarded but she looked unsure about this. Nevertheless, Amy in a rather unsure tone changed her answer that in this task they might have not discarded because she believed that the average should take into account all the data collected. She explained that an outlier can be discarded but, in this case, it cannot be done because the average is calculated.

Amy took some time to think about her answer after she was probed to explain further. She mentioned not being sure about how to handle the outlier in this situation. Amy mentioned that when a data is too far away and that the information provided by this data is not right or impossible such as watching 18 hours of TV per day might be impossible because out of 24 hours, a student spends at least 8 hours in school. Therefore, she said that this information is not correct and can be discarded.

However, for this situation Amy kept saying that the information on a group of students watched more than 12 hours of TV per day cannot be discarded because the average is for all Malaysian primary school students. Amy was not aware of any other ways to find an average except for adding all the number of hours and then divide with the number of students. The above Excerpt A4 describes Amy's point on this.

Task 5. Amy identified the median by first arranging the data in the ascending order. If the data is odd numbered, she found the median by looking at the middle number. On the hand, if the data is even numbered, she found the median by averaging the two middle numbers. Amy mentioned that the median is more adequate to represent a set of data that has an extreme value.

Table 4.88

Excerpt A5

R: Which is the weight for the median child?
S: 19 kg
R: How did you find the median?
S: You arrange the data in ascending order. Then after using these nine data, you count 1, 2, 3, 4, 5 and then count 1, 2, 3, 4, 5. {She takes the middle value}
R: Is there any other way you can obtain the median?
S: You arrange in ascending order say 9 divide by 2 then you get 4.5, so you plus 0.5 and you get the 5^{th} data. The median is the 5^{th} data.
R: Besides this is there any other ways?
S: No
R: Which is the median if we include another child who weighs 43 kg?
S: Is also 19.
R: How did you obtain that?
S: The median we obtain is between 19 and 19. So you add up this two and divide by two and you get 19.
R: Any other way to find the median?
S: You count 1 to 5 and again 1 to 5 and you get the median {she takes the average of the two middle values}

R: Is it adequate to use the mean to represent the weight of the 10 children?
S: No. For this case.
R: Why?
S: Because in this case there is 43 kg. If you want to use the mean, the data should be roughly the same.
R: How do you represent the weight of the 10 children then?
S: Median
R: Why median?
S: Because all nine data except 43 are quite close to the median. But if only dealing with these nine data, you can use the mean. But since 43 is included and because 43 is far away from the other data, so it is better to use the median.

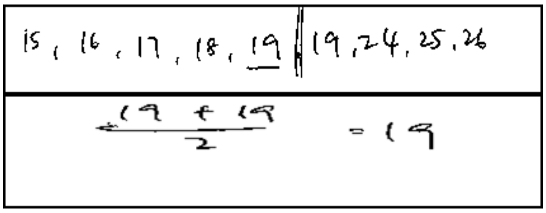


Figure 4.18. Entry A2

Amy identified the weight for the median child as 19. She explained that she identified the median by first arranging the data in ascending order. After when she counted from first to the fifth position on both sides of the arrangement to identify the middle position which is the fifth position. She took the middle value as the median.

Amy mentioned another alternative way to identify the median. She said that the total number of data which is nine is divided by two which gives four point five. She added another zero point five to four point five which gives her five. She said that this value five indicates the position of the median as the fifth position in the ascending arrangement. Amy only elaborated on these two ways to identify the median and mentioned that she is not aware of any other ways.

Amy mentioned that the median is still 19 after the inclusion of 43 kg. She explained that the median now is in between 19 and 19. Therefore she said that these two values are added up and then divided by two to give her the median as 19. This is also seen in Entry A2.

Amy mentioned that the mean is not adequate to represent the weight of all ten children. She said that this is due to the weight of the tenth child which is 43 kg. She elaborated that if the mean is considered than the data should be roughly the same. Amy mentioned that the median is more suitable because the median is almost close to nine data except 43. She added that if the data only deals with these nine data except for 43 than the mean can be used. However, she mentioned that since 43 is included in the data and since it is far away from the rest of the data, therefore she said that it is better to use the median. The above Excerpt A5 illustrates Amy's points on this.

Task 6 and task 7. Amy used the average price to guide her construction of data sets. She utilised the backward mean calculation to obtain the total price which was also used to construct the data sets. Amy preferred to play around with the prices by adding and subtracting certain values and then to come up with a few possible data sets.

Table 4.89

Excerpt A6

R: How did you place the prices?
S: All 27.
R: Why?
S: Mmm… because the average price is 27. So, you add up everything and divide by 7 is 27. So, to make things easy, I do not need to do any calculation, I put down all as 27.
R: How did you base each price?
S: Based on the average value
R: Is there any other way that you can place the average price?
S: There are a lot of ways actually. You can also place as 28, 26, 28, 26, 28, 26, and 27.
R: Why?
S: Just to make it easier because so that 28 plus 26 divide by two equals to 27 so on. There are three pairs of similar values in the construction. Therefore, I get the average as 27.
R: Any other ways?
S: Actually, there are a lot of ways. As long as you divide the total by 7, you get 27. That means any seven numbers that once you total up and then divide by 7, you get 27. Which means 27 times 7 equals to 189. Actually, the seven boxes the total numbers must add up to 189.
R: How would the price for each box?
S: It can be any value. Based on the average
R: Can you explain?
S: The average value becomes the guide so that you do not start with values like 80 and so on but values around 27

Amy placed the prices for all the bags of crisp as “All 27”. She said that the average price is 27, so when she added up all the prices and then divided with 7, the average price 27 is obtained. Amy wanted to avoid calculation and make things easy for her. Therefore, she put down all the prices as 27. Amy said that she based her decision on the average price 27. Nevertheless, Amy also mentioned that this is not the only available construction for the prices. She said that there are a lot of other ways to construct this data set.

In fact, Amy disclosed another example where she placed the prices as 28, 26, 28, 26, 28, 26, and 27. She constructed this data set by adding 28 with 26 and later divided the total by 2 which gave her 27. Amy mentioned that since there are three pairs of similar values in the data set where each pair averaged to 27, therefore the overall average for this construction will also be 27.

Amy explained that as long as the total of all the prices placed is 189 and this total after being divided with 7 gives 27, then the construction for the seven prices is accepted. She said that any value can be placed based on the average. She explained that the average price is the guide to decide on the choice of the price so that not start with big values like 80 but to start with values around 27. The above Excerpt A6 describes Amy’s points on this.

Table 4.90

Excerpt A7

R: Now look at task 7. How did you base these values?
S: I take 27 as the guide because it is the average price but none of the values can be 27. So, like just now that method 28, 26, 28, 26, 28, 26, 27, the value 27 cannot be used. I changed 27 to a different number.
R: Is there any other ways to find the values?
S: You mean any other possibilities.
R: Yes.
S: There are many other combinations actually. Just like add two or three and then minus the same on the other number. Just adjust the values by adding and subtracting.

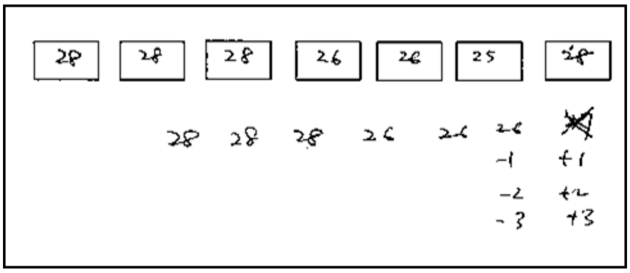


Figure 4.19. Entry A3

Similar to Task 6 in Task 7, Amy also based her construction on the average price 27. She said that she utilised 27 as her guide to decide on her choice of prices. Only that she mentioned that she would not use 27 as any of the prices. Amy used back her earlier construction 28, 26, 28, 26, 28, 26, 27. She said that 27 cannot be used and that she needed to change 27 to a different number. Amy changed 27 to another number and came up with this construction which is shown in Entry A3; 28, 28, 28, 26, 26, 25, and 28.

Amy had actually reduced 27 in the earlier construction in Task 6 to 25 and she added 2 points to one of 26s making it to 28. Amy mentioned that there are many other possible constructions. She said that to adjust the prices by adding two or three points on certain prices and subtracting the added two or three points on the other prices to give a few other possible data sets. The above Excerpt A7 illustrates this.

Summary.

Table 4.91

Subject matter knowledge with reference to context

Task	Statistical element	Conclusion
Task 1	Mean as average	Amy mentioned that the average is obtained by adding all the number of children for the families involved and then dividing the total with the number of families involved. She even elaborated this in her example. Amy also had said that the mean is the average in her responses which clearly indicated that for her the average is synonymous to the mean. She explained that when the total numbers involved were non- divisible numbers, a decimal average can be obtained.
Task 2	Idea of mode	Amy utilised the idea of mode without the task explicitly stating to do so. She said that she would place

		the order for her shop based on the frequency and the shoe size that had the highest purchase would be purchased more.
	Mode as data representation	Amy said that she would use the mode to represent the shoe sizes in her shop since the mode and the average were around the same value. However, she said if there are not around the same value then she would not use the mode to represent because she felt that the other data needed to be considered too.
	Mode as average	Amy did not use the mode as average. Instead she had calculated the average using the mean calculation. Amy said that the question is asking for the average or the mean.
Task 3	Idea of mode	Amy identified the mode based on the highest number of comments. Amy had shown that her knowledge of the mode as categorical.
	Mode as data representation	Amy said that she would not use the mode to represent the data. According to her to represent the number of comments meant that to represent for the number of comments made by all the students in the class in which she believed the mean or the average is better to be used instead of the mode.
Task 4	Median as average	Amy said that the average is obtained by adding the number of hours TV watched and then divided the total with the number of students. She even elaborated this in her example and mentioned clearly that she had carried out the mean calculation. However, regarding the extreme data, Amy said that this data would have been included in the calculation too. According to her when considering the average for all Malaysian students, the average has to consider all the data. Amy said that there are no other ways to find the average except by calculating the mean.
Task 5	Idea of median	Amy had two different ways to locate the median. She said that first the data had to be arranged in ascending order. Then she looked for the middle data. She located the middle data by counting equal number of data on both sides of the arrangement. Since the data was odd numbered, she got one middle number which was identified as the median. Amy also elaborated another way to locate the median, which is she divided the total by two and added another 0.5 to the value so that she got a whole number that indicated the position of the median in the arrangement. For even numbered data, she got the median by counting equal number of data on both sides of the arrangement. She found that there are two middle numbers in which she added both the numbers and then divided by two to give her the median.
	Robustness of measures	Amy mentioned that for this situation, the mean is not adequate to represent the data because of one extreme data. According to her, if the mean is used to represent than the data has to be roughly around the same value.

		Instead she said that the median can be used because all the other data excluding the extreme data are around the median. Amy said that if the data involved excluded the extreme, then the mean can be used.
Task 6	Idea of average	Amy placed all the values for all the seven bags of crisps as 27. She said that when all these values were added up and then divided with 7, she got the average 27. She placed all 27 to make things easy for her because she need not have to do any calculation. Amy had another construction, where she adjusted a pair of 27 to 28 and 26 in which this pair would still average out to 27. She did the same to two other pairs and remained the seventh price as 27. Amy carried out the backward mean calculation and got the total price for all the bags of crisps. She said that for any construction, the total must add up to the total price. Amy utilised the average as her guide too. She based the values which are around the average price.
Task 7	Idea of average	Similar, to the earlier task, Amy adjusted her second construction by adding two or three points on certain values and then subtracting another two or three points on the other values in order for her to come with another construction that does not carry the average price.

Subject matter knowledge of measures of central tendency in handling bias.

Task 8 and task 9. Amy chose the median as the best measure to represent the data. Amy mentioned that if the mean was calculated after the extreme data was excluded, then this mean can be considered to represent the data. However, she was not consistent with this knowledge.

Table 4.92

Excerpt A8

R: What are the measures of central tendency that you can obtain based on the above data?
S: Mode and median
R: What is the median?
S: 6.2
R: What is the mode?
S: 6.3
R: How did you find the median?
S: There are 9 data. So, arrange ascending order and the fifth data will be the median.
R: Any other way that you can find the median?
S: Mmm…no
R: How about the mode?
S: The highest number of frequency that occurs. 6.3 got the highest frequency, so 6.3 is the mode.
R: How about the mean?

S: Mean can be calculated but it is not a good measure to use here because there is a value which is far apart from the other values.
R: But can you obtain the mean?
S: Yes
R: How?
S: Add all this up (points at the data in the task sheet) and divide by nine.
R: Which is the best measure of central tendency to represent the actual weight of this object?
S: Median
R: Why?
S: Mean is not suitable because there is 15.3. Median is the best because all the data are arranged in ascending and the middle value is taken as the median. But mode only takes care of the group with the highest frequency. That means it doesn't take count the rest of the values.
R: How about the mean?
S: [Paused a while] Eliminate the 15.3. Because they are weighing the same object. So, all the data here is around 6.0 to 6.3. Probably this one 15.3 is an error which can be eliminated.
R: So how?
S: If this data is not included then mean can be considered. {Calculated the mean as 6.17 using the calculator}
R: Now, which is the best measure to represent the actual weight of this object?
S: Median. Maybe mean can be considered if calculated after eliminating 15.3
R: Why?
S: Because the median 6.2 and the mean 6.17 are around the same value as the other data. These two measures consider all the data.

Amy mentioned that for the data in this task, the median is 6.2 and the mode is 6.3. She explained that the median was found after all the nine data were arranged in ascending order and the fifth data which is the middle data was taken as the median. She said that this is the only way to find the median. Amy said that the mode was found based on the highest number of frequency that occurs. She said that 6.3 got the highest frequency, therefore 6.3 is the mode.

Amy claimed that the mean can be calculated but it is not a good measure to be used in this situation because of one particular value which is far apart from the other values in the data. However, she mentioned that the mean can be obtained by adding all the values in the data and then divide the total with nine.

Amy mentioned that the best measure to represent the actual weight of this object is the median. She said that the mean is not suitable because of 15.3. According to her, the median is the best because all the data are arranged in ascending order first and the middle value in the arrangement is taken as the median. On the other hand, she felt that the mode only considers the data with the highest frequency and does not take count the rest of the values.

Although Amy mentioned earlier that the mean is not suitable because of 15.3. She mentioned that 15.3 looked more like a measurement error because the rest of the data is around 6.0 to 6.3. She added that if 15.3 is eliminated than the mean can be considered. Amy mentioned that not only the median can best represent the actual weight of this object but maybe the mean can be considered also if it is recalculated after 15.3 is eliminated. She explained that this is because the median which is 6.2 and the mean which was calculated after 15.3 was eliminated is 6.17 are around the same value with the rest of the data. The above Excerpt A8 describes Amy's points on this.

Table 4.93

Excerpt A9

R: What are the measures of central tendency that you can obtain based on the above data?
S: The median and the mean
R: How did you find the median?
S: Arrange the data from the lowest to highest value. There are five data, so you take the third data which is 54 000.
R: Any other way that you can find the median?
S: No
R: How about the mean?
S: Mean, you add all up and you divide by 5.
R: What about the mode?
S: There is no mode here.
R: Why?
S: Because all the data are of different values. No highest frequency here.
R: Which is the best measure of central tendency to represent the actual salary of the company?
S: Median
R: Why?
S: Because employee E has a salary which is very high so that means this data is very different from the rest of the data. So, it is better to use median
R: How about the mean?
S: Cannot because the mean is not accurate.
R: Why?
S: Because it is salary of the company. It must consider all the data. In this case, the mean which I calculated as 76 800 is high and does not represent the most salaries.

Amy mentioned that only the median and the mean can be obtained based on the data given in Task 9. She said that she found the median as 54 000 which is the third data after all the five data were arranged in the ascending order. Similar to the earlier task, she said that this is the only way to find the median. As for the mean, she said that added up all the five data and then divided the total with five. She also mentioned that there is no mode in this data because all the data were of different values and there is no highest frequency.

According to Amy, the median is the best measure to represent the actual salary of the company. She explained that employee E had a very high salary in which is very different than the rest of the data, therefore she felt that the median is the best measure to represent the actual salary of the company. On the other hand, Amy said that the mean cannot represent the actual salary of the company because she claimed that it is not accurate. Amy explained that since the data is related to a company's salary, all data has to be considered in the mean calculation. She said that the mean 76 800 that she calculated was high and does not represent the majority of the company's salaries. The above Excerpt A9 illustrates Amy's points on this.

Summary.

Table 4.94

Subject matter knowledge in handling bias

Task	Statistical element	Conclusion
Task 8	Identifies and summarises data using mean	Amy identified and calculated the mean by adding all the given data and then divided the total with the total number of data.
	Identifies and summarises data using median	Amy identified and obtained the median by first arranging the data in ascending order and then she looked for the middle value because the data was odd numbered. This middle value gave her the median. Amy said that there are no other ways to locate the median.
	Identifies and summarises data using mode	Amy identified the mode based on the highest frequency.
	Best data representation measure	Amy mentioned that the median is the best representation measure. She said this is because the median is found after arranging all the data in ascending order and the middle value is taken as the median. She also said that the median value is around the values of the other data except the extreme. Amy said that the mean is not suitable because it included the extreme data whereas the mode is based on the data with the highest frequency and does not consider the rest of the data. However, Amy said that the mean can be considered if the extreme data was excluded from the

		mean calculation. This mean would now be around the value of most of the data.
Task 9	Identifies and summarises data using mean	Amy identified and calculated the mean by adding all the given data and then divided the total with the total number of data.
	Identifies and summarises data using median	Amy identified and obtained the median by first arranging the data in ascending order and then she looked for the middle value because the data was odd numbered. This middle value gave her the median. Amy said that there are no other ways to locate the median.
	Identifies and summarises data using mode	Amy mentioned that there is no mode for this data because there are no data with the highest frequency.
	Best data representation measure	Amy mentioned that the median is the best to represent the salary because of the extreme data. According to her, the mean cannot be used here because she said it is not accurate. She explained that since the data involved the salaries of a company so all the data need to be considered in the mean calculation. She found that when all the data was considered the mean does not represent the most salaries.

Subject matter knowledge of measures of central tendency in problem solving.

Task 10. Initially Amy found the overall GPA by adding the GPAs of both colleges and then divided the total by two. But later she realised that there were also number of semesters involved in colleges which made her to calculate the overall GPA differently.

Table 4.95

Excerpt A10

R: How do you find the overall GPA?
S: Because college A got 3.2 and then for college B is 3.8. So, you add up the GPA for college A with GPA for college B. 3.2 plus 3.8 and then divide by 2. You get 3.5. [Paused for a while and was thinking] So for two semesters does it mean that in college A, the first semester he got 3.2 and the second semester he got 3.2 also? It is like this. So, 3.2 times 2 plus 3.8 times 3. (Writes down the answer)
R: Why?
S: Because we got cumulative of 3.2 for two semesters so we times two. For 3.8 for three semesters so we times 3.8 with three. Then add up the total and then divide with five semesters. You get 3.56
R: Why divide with five?
S: Because we take account of five semesters
R: Earlier you mentioned the overall GPA is 3.5 and then you said 3.56. Which one is it?
S: The second one 3.56 is the correct overall GPA. The earlier one, I thought that is calculated as for two colleges that means as GPA for college A and college B. I did not see the semesters clearly.

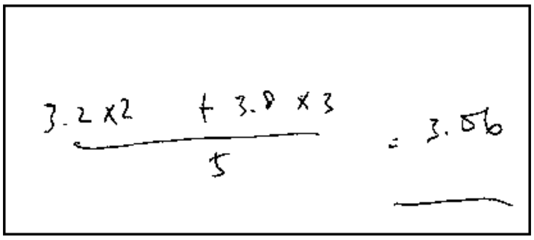


Figure 4.20. Entry A4

Initially Amy mentioned that she found the overall GPA by adding the GPA for college A 3.2 with the GPA for college B 3.8. Then she divided the total with two and got the overall GPA as 3.5. However, after calculating this overall GPA, Amy paused for a while and thought about it again while she reread the question. She realized that the GPA for college A was for two semesters. She repeated that for college A, the student got 3.2 for the first semester and another 3.2 for the second semester. After which Amy explained that the overall GPA calculation should be 3.2 multiplied with two plus 3.8 multiplied with three. She added the total grade points then divided it with five to get the overall GPA as 3.56. She wrote this clearly in the task sheet as shown in Entry A4.

Amy explained that the total grade point is divided with five because to take count for the five semesters. Amy clarified that the correct overall GPA is 3.56. She explained that the earlier GPA which is 3.5 was calculated based on the two colleges. She claimed that she did not see the semesters involved clearly which she only realized later. Excerpt A10 describes Amy's points on this.

Task 11. Amy calculated the weighted mean when she was asked to find the average. She explained that the average given in each of the groups is the group average and based on this, she calculated the total TV viewing time per weekend for each group. She utilised this total to calculate the average TV viewing time per weekend for all 100 students.

Table 4.96

Excerpt A11

R: How did you find the average?
S: It says that 25 rural students watch 8 hours, so find the total for rural students which is 25 times 8 and then find for the urban students 75 urban students watch 4 hours, so the total is 75 times 4. This is for the rural and this is for the urban (explains while referring to the answer written in the task sheet). And then you total and divide with 100 to get the average.
R: Why divide by 100?
S: Because there are 100 students overall.
R: Why are you totalling up this two? (Referring to the total of 25 times 8 and 75 times 4)
S: Because if you want to calculate the average you need to total up the number of number of hours that all the students watch and then divide by the number of students to find how many hours each student watch.

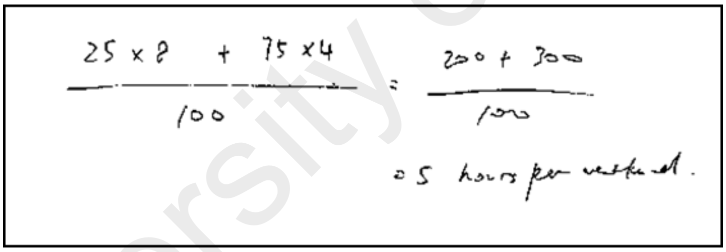


Figure 4.21. Entry A5

Amy mentioned that 25 rural students watched an average of 8 hours of TV. Therefore, she said that the total number of hours of TV watched by rural students is found by multiplying 25 with 8. Similarly, she explained that 75 urban students watched an average of 4 hours of TV, therefore the total number of hours of TV watched by urban students is 75 times 4. After when Amy mentioned that she added both the total for rural and urban and divided the overall total with 100. Amy wrote this clearly in the task sheet as shown in Entry A5.

Amy explained that the total number of hours of TV watched by rural and urban students are added up because the average is for the overall students and this total is divided with 100 because of the very same reason which is to find the average for all 100 students. The above Excerpt A11 describes this.

Summary.

Table 4.97

Subject matter knowledge in problem solving

Task	Statistical element	Conclusion
Task 10	Backward mean calculation	Amy solved for the weighted mean for both the tasks without the task explicitly stating to do so. She had carried out the backward mean calculation, where she obtained the total for each group of data set from the given group means.
	Representative nature of the mean	Amy used the knowledge related to the representative nature of the mean in making sense of why data sets need to be combined in the weighted mean.
	Forward mean calculation	Amy also carried out the "forward" mean calculation to calculate the weighted mean.
Task 11	Backward mean calculation	Amy solved for the weighted mean for both the tasks without the task explicitly stating to do so. She had carried out the "backward" mean calculation, where she obtained the total for each group of data set from the given group means.
	Representative nature of the mean	Amy used the knowledge related to the representative nature of the mean in making sense of why data sets need to be combined in the weighted mean.
	Forward mean calculation	Amy also carried out the "forward" mean calculation to calculate the weighted mean.

Subject matter knowledge of measures of central tendency in making inference.

Task 12. Amy recommended the player for All Star team by comparing the total scores. She said that the average can also be used because it is the average score per game.

Table 4.98

Excerpt A12

R: If coach Ahmad can only select one of the two players, who would you recommend he select?
S: I recommend Beng
R: Why?
S: If you add up Beng's scores is higher than Ramu's.
R: Can you explain?
R: So, you based your answer based on the total score.
S: I based on the total scores that I added all the scores of both players. But if you want to calculate average also can because average is like per game and because they played six games so the average for one. Beng's average is 21.83 and Ramu's average is 19.67.
R: Who would you recommend?
S: Beng
R: How would you select the player?
S: I can compare the total scores or average. Beng got higher total scores or average.
R: Is there any other ways to determine the player?
S: Mmm…no.

Amy recommended Beng to be selected. She said that she based the selection on the total scores in which she added all the scores for both the players. She found that Beng's total scores was higher than Ramu's. She also mentioned that the average score can be calculated because it gives the average score per game. She added that since both the players played a set of six games, therefore the average was for one game. Amy compared the averages for both the players. She found that Beng's average was 21.83 better than Ramu's average which was 19.67, therefore Beng was recommended. Amy concluded that both the total scores and the average comparison can be used to select the player. Beng had higher total scores and average, therefore she said he was recommended. Amy claimed that she is not aware of any other ways to select the players except the ones that she mentioned earlier. Excerpt A12 describes Amy's points on this.

Task 13. Amy mentioned that for unequal sized data, she would not use the total scores comparison to select the player. Instead she said that she would recommend based on the average score comparison.

Table 4.99

Excerpt A13

R: Who would he select from the group below?
S: Khan
R: Why?
S: Because Khan has higher average score compared to Ali.
R: Is there any other ways to select the player?
S: No other ways. I use the average. I cannot use the total scores here.
R: Why?
S: Because Khan and Ali played different number of games. To be fair Ali played 8 games so divide the total scores by 8. Khan played 6 times and to be fair divide the total scores by 6.
R: What is that you calculated?
S: The average score for Khan and Ali.
R: Is there any other way to choose the best player from the group?
S: [paused for a while] no.

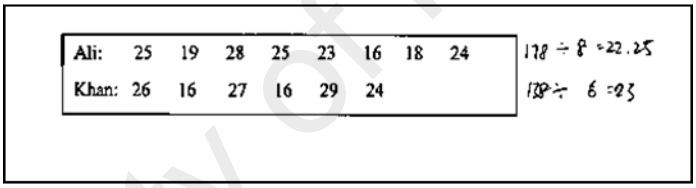


Figure 4.22. Entry A6

Amy recommended Khan to be selected because Khan had higher average score compared to Ali. This is shown in Entry A6. She said that there are no other ways except using the average comparison to select the players. She mentioned that the total scores comparison cannot be used here because both players played different number of matches. Amy mentioned that to be fair, Ali who played eight games, his total scores should be divided by eight and Khan who played six games, his total scores should be divided by six. She said she had actually calculated the average score for Khan and Ali. The above Excerpt A13 describes Amy’s points on this.

Task 14. Amy compared the mean for the number of questions correct to decide on the class that performed better.

Table 4.100

Excerpt A14

R: Based on the graphs below, which class did better?
S: Class 5A did better.
R: Why?
S: Because I calculated through this way where for 3 questions correct got 1 student, 4 questions correct got 2 students, 5 questions correct got 3 students and 6 questions correct got 3 students. I times the number of questions correct with the number of students and totalled up the values which I got 44. Then I divided the total with 9 students. I got 4.89.
R: What is this value?
S: Average. So, for class 5B, I used the same method also. For 3 questions correct 2 students, 4 questions 4 students, 5 questions correct 2 students and 6 questions correct 1 student. I times the number of questions correct with the number of students and then I add up all. The total number of questions correct, I then divided with 9 students. The average for 5B is 4.22.
R: So which class did better?
S: The one with the higher average value. 5A did better.

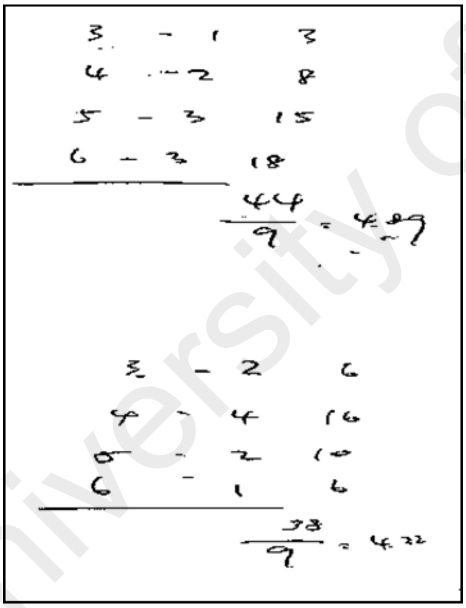


Figure 4.23. Entry A7

Amy mentioned that 5A performed better. She based her decision on the average value that she calculated. Amy elaborated that the average was calculated by first multiplying the number of questions correct with the respective number of students. After which she totalled up all the products from the earlier multiplications and then divided it with the total number of students involved.

Amy elaborated this in the example. She said that she calculated by first multiplying 3 questions correct with 1 student, 4 questions correct with 2 students, 5 questions correct with 3 students and 6 questions correct with 3 students. She totalled up the products and she said that she got 44. After which she divided the total with 9 students and got the average for 5A as 4.89.

Similar calculation was carried out with 5B. She said that she multiplied 3 questions correct with 2 students, 4 questions correct with 4 students, 5 questions correct with 2 students and 6 questions correct with 1 student. She added up all the products and then divided the total with 9 students. She got the average for 5B as 4.22. This is shown clearly in Entry A7.

Amy compared both the averages for 5A and 5B. She found that the average for 5A was higher than 5B in which she concluded that 5A performed better. The above Excerpt A14 illustrates Amy's points on this.

Task 15. Amy compared the class average to decide on the class that performed better. She also compared the shape of the graphs and the mode to decide on the class that performed better. However, Amy concluded that since both classes had different number of students, therefore the average comparison is better to be used to decide on the class performance.

Table 4.101

Excerpt A15

R: Based on the graphs below, which class did better?
S: Class 5D did better
R: Why?
S: According to the average score. I multiply the number of questions correct with the number of students. 2 questions correct 1 student, 3 questions correct 4 students, 4 questions correct 6 students and so…..(points at the working on the task sheet). I add all these up and then I divide with the total number of students. Here for 5C 198 divide 36, I get 5.5. The same goes to 5D. (Points at the working in the task sheet). I divide 130 with 21 and get the average as 6.19. The average for 5D is higher than 5C, so 5D did better
R: Is there any other way that you can use to decide which class did better?
S: You can see that for this graph, the peak is actually here (5C) so is only 5 or 6 but this one (5D) the peak is actually at 7.
R: What does that tell you?

S: You can say that this is actually the mode. When there are, more students got 7 corrects (5D) and for this class (5C) is only 5 or 6. That class is better.
R: Can you explain further?
S: The shape of the graph. Because here the higher number of corrects and here is the lower number of corrects so if the graph is this side you can know that more students got more number of corrects (5D). So, this side that means, students scored better (points at 5D) compared to this graph (points at 5C). If according to the graph if the peak is here that means that the mode is higher at this area that means more students scored more number of corrects (5D) as compared to this one (5C). But it is also difficult to interpret from the graph because the number of students involved not same. I think it is better to use average.

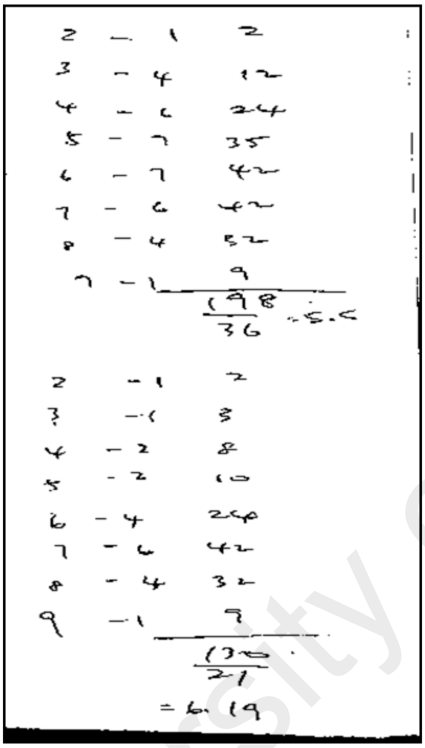


Figure 4.24. Entry A8

Amy mentioned that 5D performed better. She based this on the average. She said that she calculated by first multiplying the number of questions correct with the respective number of students. Then she totalled up all the products and divided the total with the total number of students in the class. For 5C, Amy divided 198 with 36 and got the average as 5.5. For 5D, she divided 130 with 21 and got the average as 6.19. She mentioned that since the average for 5D is higher than 5C, thus 5D did better. This is shown in Amy's working in Entry A8.

According to Amy, another way to decide on the class performance would be by looking at the graphs. She mentioned that the peak for the graph 5C is somewhere around number of corrects 5 to 6 whereas the peak for the graph 5D is somewhere around number of corrects 7. She said that the peak is where the mode is located for both the classes. According to her, when there are more students who got more number of corrects, in this case more students in 5D got 7 corrects as compared to 5C that only got 5 or 6 corrects, therefore class 5D performed better than 5C.

Amy elaborated that from the shape of the graphs and by looking at the parts of the graphs that show the number of corrects, 5D's graph was inclined towards more number of corrects. The students in 5D scored better than 5C. 5D also had the peak at the higher number of corrects and the mode for 5D is also higher than 5C. Thus, Amy mentioned that more students in 5D scored better. Nevertheless, Amy finally concluded that this graph comparison method is difficult to be used because the number of students for both classes is different. She believed that the average comparison is still better to decide on the class performance. The above Excerpt A15 illustrates this.

Summary.

Table 4.102

Subject matter knowledge in making inference

Task	Statistical element	Conclusion
Task 12	Summarises equal sized numerical data using measures of central tendency	Amy summarised equal sized numerical data using the mean.
	Utilises the appropriate measure to compare and draw conclusion about equal sized numerical data	Amy compared two equal sized numerical data using the total scores and average and decided to recommend the player to be chosen based on this.
Task 13	Summarises unequal sized numerical data using measures of central tendency	Amy summarised two unequal sized numerical data using the mean.

	Utilises the appropriate measure to compare and draw conclusion about unequal sized numerical data	Amy compared two unequal sized numerical data using the average and decided to recommend the player to be chosen based on this. She said that here she would not use the total scores because both the players played different number of sets. According to her, it is only fair to divide the total scores with the total games played by each player.
Task 14	Summarises equal sized graphical data using measures of central tendency	Amy summarised equal sized graphical data using the average.
	Utilises the appropriate measure to compare and draw conclusion about equal sized graphical data	Amy compared two equal sized graphical data using the average and decided on the player to be selected.
Task 15	Summarises unequal sized graphical data using measures of central tendency	Amy summarised two unequal sized graphical data using the mean and the mode.
	Utilises the appropriate measure to compare and draw conclusion about unequal sized graphical data	Amy had several ways to compare and decide on the player to be selected. She compared the average. She also compared the mode. Amy also looked at the shape of the graphs by seeing which graph had more students with more number of questions correct. However, among all the ways that she mentioned, Amy said that the average is better to be used in the selection of the player to be recommended.

Level of subject matter knowledge of measures of central tendency.

Level of subject matter knowledge of measures of central tendency with reference to context. In Task 1, Amy mentioned that the average was found by adding up all the children from all the families and then divided the total with the number of families. She also showed the mean calculation in her example. Amy mentioned that when two non-divisible numbers are involved in the calculation of the average, a decimal average is possible. She also mentioned that the average here referred to the mean. Therefore, Amy had presented the knowledge of the mean as the average (P) and her justification was found appropriate too (A).

In Task 2, Amy had utilised the idea of mode without the task explicitly stating so. She found the frequency for each shoe size and had identified the mode as 5 based on the highest purchase or highest frequency. Therefore, Amy's idea of the mode was found to be present (P) and correct (A).

Amy said that the mode can be a form of data representation if the mode and the mean were around the same values. Since the mode was 5 and the mean was 5.5, so she said that the mode can be used to represent the data here. However, she said that the mode cannot represent the data if both the mean and the mode values are not around the same values because she said that the other data also play an important part in the data representation. Therefore, Amy's knowledge of the mode as a form of data representation was marked as not present (NP) because the mode can be utilised as a form of data representation regardless the mean value and the absence of this knowledge is considered problematic (I).

Amy did not utilise the mode as the average. Instead she calculated the average as 5.5. According to her, the average means the mean and she calculated the average using the mean algorithm. Amy had lacked in the knowledge that the mode can be used as a quick method in the reporting of an average. Therefore, Amy's knowledge of the mode as the average was not marked not present (NP) and the absence was considered problematic (I).

In Task 3, Amy identified the mode as 22. She took the mode as the highest number of comments. When the task was modified where the number of comments for student D was changed from 22 to 6, Amy identified the mode as 6 in which she still referred to the highest number of comments. Amy had identified the mode incorrectly. First, she had taken the highest number of comments to be the mode in which she believed to be highest frequency. Amy also lacked in the knowledge that the mode can be categorical. The mode is referred to be the category that carried the highest frequency. Therefore, Amy's idea of the mode was marked to be not present (NP) and the lack is considered inappropriate (I).

Amy mentioned that she would not utilise the mode to represent the number of comments made by the students. She said that the mode was found to be extreme than the other values in the data (for the case where the number of comments for student D was 22). Even she still believed that the mode cannot be utilised to represent the number of comments for the case where the number of comments for student D was changed to 6.

Instead she said that it is better to use the average. According to her, to represent the number of comments it is better to use the number of comments made by all the students in the class, so the mean is better. Therefore, Amy's knowledge of the mode as a form of data representation was found to be not present (NP) and the lack was considered problematic (I).

In Task 4, Amy mentioned that the average is by adding the number of hours students watched TV and then divided the total with the number of students. She even elaborated this in her example and mentioned that she was actually showing the mean calculation. However, Amy was unsure about the information on extreme data. First she said that they would have add it into the mean calculation. Then she said that this information would have been discarded. Amy knew that an outlier can be discarded but she claimed that the outlier cannot be discarded in calculating the average. She added that the average here was for one whole Malaysia and all the information should be considered including the extreme data.

Normally an average involving data on human population would be based on the median because the median is least likely to be affected by the extreme data. However, Amy did not present the knowledge of the median as the average here (NP). She said that the average was based on the mean which is acceptable but given to proper justification. However, Amy's said that the mean was calculated based on all the data including the extreme which is considered inappropriate (I) because the mean calculated after excluding the extreme would have been better.

In Task 5, Amy found the median as 19. She found the median by rearranging all the weights in ascending order. For the data involving nine children, she identified the middle or the fifth data as the median. Amy utilised two different ways to locate the middle data, one she counted both sides of the arrangement and located the middle value. Second, she divided the total number of data with 2, she got 4.5 in which she added another 0.5 to round off the value to 5. She located the 5^{th} data as the middle data.

Amy said that the median is still 19 after 43 kg was included in the data. She said that the median is between 19 and 19, in which she added both these values and then divided the total by two. Thus, Amy's knowledge of the idea of the median was found to be present (P) and appropriate because she carried and identified the median correctly (A).

Amy said that the mean is not adequate to represent the weight of all ten children. She explained that this was because the weight of the tenth child was much heavier compared to the rest. Instead Amy said to consider the median to represent the weight of the children. She said that the mean is not suitable because the mean should be considered when the data are roughly the same.

She explained that nine out of the ten data are quite close to the median. However, Amy also mentioned that if only dealing with the nine children, then the mean can be considered but since 43 kg is also included than the median is better to be used. Thus, Amy's knowledge related to the robustness of the measures of central tendency was found to be present (P) and appropriate (A) because she could choose and justify the appropriate measure to represent the data.

In Task 6 and Task 7, Amy constructed data sets based on the given average value and the conditions imposed. In Task 6, Amy put down the prices as all 27. She said when these values are added up and then divided with 7, she gets 27. So, it was easy for her to put all 27. She also said that there are the ways too. Another example, she included the pair average 27 meaning a pair 28 and 26 averages to 27. She had three sets of these pair and the remaining seventh value was placed as 27.

Amy utilised the backward mean calculation and got the total price as 189. She said that the prices can be of any values but when totalled up must be 189. She also said that the average value became a guide for her to choose values that are around 27. Therefore, Amy's knowledge of constructing a data set based on the idea of average was marked present (P) and found to be appropriate too (A).

In Task 7, Amy placed the prices in a similar way as how she did in Task 6. She calculated the total price. She used the average as the guide. She adjusted the values in the earlier construction by adding and subtracting two to three numbers to get other possible combinations. Amy's knowledge of constructing a data set based on the idea of average for Task 7 was found to be present (P) and appropriate (A).

Table 4.103

Coding rubrics for determining level of Amy's subject matter knowledge of measures of central tendency with reference to context

Task	Subject Matter Knowledge of Measures of Central Tendency with reference to Context																															
	Mean as average				Mode as average				Median as average				Idea of mode				Idea of median				Mode as data representation				Robustness of measures				Idea of average			
	P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP	
	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I
Task 1	X																															
Task 2								X					X											X								
Task 3																X								X								
Task 4												X																				
Task 5																	X								X							
Task 6																													X			
Task 7																													X			
	1												**1**				**1**								**1**				**2**			

Legend: P = Present NP = Not Present A = Appropriate I = Inappropriate

Amy's Percentage of Subject Matter Knowledge of Measures of Central Tendency with Reference to Context

$$= \frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)} \text{ X } 100\% = \frac{\mathbf{6}}{\mathbf{11}} \times \mathbf{100\%} = 54.55\% \text{ (Medium)}$$

Level of subject matter knowledge of measures of central tendency in making inference. In Task 8, Amy calculated the mean for all nine data. She added up all the nine data and then divided the total with nine. Therefore, Amy not only identified that the mean can be obtained but also summarised the data using the mean by calculating the mean value. Therefore, Amy's knowledge of identifying and summarising the data using the mean was found to be present (P) and correct (A).

Amy said that the median can be found after the data was arranged in ascending order and the fifth data which was 6.2 as the median. Thus, Amy's knowledge of identifying and summarising the data using the median was found to be present (P) and correct (A).

Amy identified the mode was 6.3 based on the highest frequency. Therefore, Amy's knowledge of identifying and summarising the data using the mode was found to be present (P) and correct (A).

Amy mentioned that the median was the best measure to represent the actual weight of this object. She said that the mean is not suitable because of 15.3. However, she said that the mean can be considered if 15.3 was eliminated from the data. She calculated the mean after excluding 15.3 and got the value as 6.17. Amy mentioned that the median 6.2 and the mean 6.17 were around the values of the other data and these two measures considered all data, therefore can represent the data. Thus. Amy's knowledge related to the best representation measure was found to be present (P) and her justifications were also appropriate (A)

Similarly, in Task 9, Amy calculated the mean for all five data. She added up all the five data and then divided the total with five. She got the mean as 76 800. Amy not only identified that the mean can be obtained but also summarised the data using the mean by calculating the mean value. Therefore, Amy's knowledge of identifying and summarising the data using the mean was found to be present (P) and correct (A).

Amy found the median as 48 000. She arranged the data in the ascending order and identified the third data as the median. Thus, Amy's knowledge of identifying and summarising the data using the median was found to be present (P) and correct (A).

Amy mentioned that there was no mode for the given data because there was no data with the highest frequency. Therefore, Amy's knowledge of identifying and summarising the data using the mode was found to be present (P) and correct (A).

Amy mentioned that the median was the best measure to represent the data because of the extreme data involving the salary of employee E. She said that the mean cannot be used here because the mean was not accurate. According to her, the mean involving the salary of a company has to consider all the salaries. She said that since the mean was high, therefore it does not represent the most salaries. Thus, Amy's knowledge related to the best representation measure was found to be present (P) and appropriate given to her justification (A).

Table 4.104

Coding rubrics for determining level of Amy's subject matter knowledge of measures of central tendency in handling bias

Task	Subject Matter Knowledge of Measures of Central Tendency in Handling Bias															
	Identifies and summarises data using mean				Identifies and summarises data using median				Identifies and summarises data using mode				Best data representation measure			
	P		NP		P		NP		P		NP		P		NP	
	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I
8	X				X				X				X			
9	X				X				X				X			
	2				**2**				**2**				**2**			

Legend: P = Present NP = Not Present A = Appropriate I = Inappropriate

Amy's Percentage of Subject Matter Knowledge of Measures of Central Tendency in Handling Bias

$$= \frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)} \text{ X } 100\% = \frac{\mathbf{8}}{\mathbf{8}} \times \mathbf{100\%} = 100\% \text{ (High)}$$

Level of subject matter knowledge of measures of central tendency in problem solving. In Task 10, Amy carried out the backward mean calculation to obtain the total grade point average for College A based on two semesters when she multiplied 3.2 with 2. She also found the total grade point average for College B based on three semesters when she multiplied 3.8 with 3. Therefore, Amy's knowledge of performing the backward mean calculation to obtain the total from the mean is found to present (P) and correct (A).

Amy explained that the mean which was given for each college was for the number of semesters involved in that particular college. She also added the total grade points of both colleges and the total semesters of both colleges before she proceeded in the weighted mean calculation. Thus, Amy's knowledge related to the representative nature of the mean is found to be present (P) and appropriate (A).

Finally, Amy carried out the forward mean calculation to find the weighted mean. She divided the total grade points of both colleges with the total semesters of both colleges and found the weighted mean as 3.56. Therefore, Amy's knowledge of the forward mean calculation is found to present (P) and appropriate (A) despite the task not explicitly stating to find for the weighted mean.

In Task 11, Amy performed the backward mean calculation and obtained the total hours of TV watched per weekend for both rural and urban groups. She said that she multiplied the number of students in each group with the group average. Amy found the total hours of TV watched for rural group by multiplying 25 with 8. Similarly, for the urban group she multiplied 75 with 4. Thus, Amy's knowledge of performing backward mean calculation was found to be present (P) and appropriate (A).

Amy explained that she multiplied the average of each group with number of students in each group because the given average was for the group. She also mentioned that the totals of both the groups were added and the total number of students of both groups were found before she proceeded with further calculation because she wanted to find the overall average for all the students. Thus, Amy's knowledge related to the representative nature of the mean is found to be present (P) and appropriate (A).

Finally, Amy carried out the forward mean calculation to find the weighted mean. She divided the total number of hours of both groups with the total number of students and found the average as 5. Therefore, Amy's knowledge of the forward mean calculation is found to present (P) and appropriate (A) despite the task not explicitly stating to find for the weighted mean.

Table 4.105

Coding rubrics for determining level of Amy's subject matter knowledge of measures of central tendency in problem solving

Task	Subject Matter Knowledge of Measures of Central Tendency in Problem Solving											
	Backward mean calculation				Representative nature of the mean				Forward mean calculation			
	P		NP		P		NP		P		NP	
	A	I	A	I	A	I	A	I	A	I	A	I
10	X				X				X			
11	X				X				X			
	2				**2**				**2**			

Legend: P = Present NP = Not Present A = Appropriate I = Inappropriate

Amy's Percentage of Subject Matter Knowledge of Measures of Central Tendency in Problem Solving

$$= \frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)} \text{ X } 100\% = \frac{\mathbf{6}}{\mathbf{6}} \times \mathbf{100}\% = 100\% \text{ (High)}$$

Level of subject matter knowledge of measures of central tendency in making inference. In Task 12, Amy calculated the mean scores for both the players by adding all the scores and divided the total with the six set of games that they played. She found that the mean score for Ramu was 19.67 and the mean score for Beng was 21.83. Therefore, Amy's knowledge of summarising equal sized numerical data using the measures of central tendency in this case the mean without the task explicitly stating to do so was found to be present (P) and appropriate (A).

Amy compared the mean scores of both the players and decided on the player that she would recommend to be selected for the team. Since Beng's average was higher than Ramu's, she said that Beng would be recommended. Amy also mentioned that the player can also be selected by comparing their total scores. Thus, Amy's knowledge of utilising the appropriate measure of central tendency to compare and draw conclusion on two equal sized numerical data was found to be present (P) and appropriate (A)

In Task 13, Amy summarised the data using the mean or the average. She calculated the average by adding all the scores for each player and then divided the total scores with the number of game sets that they played. She said for Ali, she divided his total scores with eight because he played eight game sets. Whereas for Khan, she divided his total scores with six because he played six game sets. Amy found that Ali's average was 22.25 while Khan's average was 23. Therefore, Amy's knowledge of summarising unequal sized numerical data using the measures of central tendency in this case the mean without the task explicitly stating to do so was found to be present (P) and correct (A).

Amy compared the average scores of both the players and decided to recommend Khan because his average score was higher than Ali's. Amy said that the average was a fair measure because all the game sets that the player played were taken count. Amy said that since the number of game sets played by the players were different, therefore the total score comparison cannot be done here.

Thus, Amy's knowledge of utilising the appropriate measure of central tendency in this case the mean to compare and draw conclusion on two unequal sized numerical data was found to be present (P) because the mean is the most suitable measure to be used to compare and draw conclusions based on unequal sized data. Her knowledge of utilising the mean here was considered to be accurate too (A).

In Task 14, Amy summarised the data given using the mean or the average. She multiplied the number of students with the number of corrects and added the products. Then she divided the total with the total number of students. Amy calculated the average for 5A as 4.89 and for 5B as 4.22. Therefore, Amy's knowledge of summarising equal sized graphical data using the measures of central tendency in this case the mean without the task explicitly stating to do so was found to be present (P) and her knowledge of calculating the mean was found to be appropriate (A).

Amy compared the class average of both the classes and found that 5A had a higher average than 5B. Therefore, she said that 5A performed better than 5B. Thus, Amy's knowledge of utilising the appropriate measure in this case the mean to compare and draw conclusion on two equal sized graphical data was found to be present (P). She also had calculated the average or the mean correctly and explained that the class that had higher average performed better, so her justification here is considered appropriate (A).

In Task 15, Amy summarised the data given using the mean and the mode. She multiplied the number of students with the number of corrects and added the products. Then she divided the total with the total number of students. Amy got the average for

5C as 5.5 and for 5D as 6.19. She also found that the mode for 5C is 5 or 6 and for 5D is 7. Therefore, Amy's knowledge of summarising unequal sized graphical data using the measures of central tendency in this case the mean and the mode were used without the task explicitly stating to do so was found to be present (P) and her knowledge of calculating the mean and identifying the mode were found to be appropriate too (A).

Amy compared the average of both the classes and decided that 5D did better because 5D had a higher average than 5C. She also compared the mode and the graphs and based on this she said that 5D did better. However, due to the unequal number of students involved for both the classes, Amy said that the average comparison is the best to determine the class that performed better. Thus, Amy's knowledge of utilising the appropriate measure in this case the mean to compare and draw conclusion on two unequal sized graphical data was found to be present (P). She also had calculated the average or the mean correctly and explained that the class that had higher average performed better, so her justification here is considered appropriate too (A).

Table 4.106

Coding rubrics for determining level of Amy's subject matter knowledge of measures of central tendency in making inference

	Subject Matter Knowledge of Measures of Central Tendency in Making Inference																															
Task	Summarises two equal sized numerical data using measures of central tendency				Utilises the appropriate measure to compare and draw conclusion about two equal sized numerical data				Summarises two unequal sized numerical data using measures of central tendency				Utilises the appropriate measure to compare and draw conclusion about two unequal sized numerical data				Summarises two equal sized graphical data using measures of central tendency				Utilises the appropriate measure to compare and draw conclusion about two equal sized graphical data				Summarises two unequal sized graphical data using measures of central tendency				Utilises the appropriate measure to compare and draw conclusion about two unequal sized graphical data			
	P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP	
	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I
Task 12	X				X																											
Task 13									X				X																			
Task 14																	X				X											
Task 15																									X				X			
	1				**1**				**1**				**1**				**1**				**1**				**1**				**1**			

Legend: P = Present NP = Not Present A = Appropriate I = Inappropriate

Amy's Percentage of Subject Matter Knowledge of Measures of Central Tendency in Making Inference

$$= \frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)} \text{ X } 100\% = \frac{\mathbf{8}}{\mathbf{8}} \times \mathbf{100\%} = 100\% \text{ (High)}$$

Amy's mean percentage of subject matter knowledge for all four constructs $\frac{\mathbf{54.55+100+100+100}}{\mathbf{4}} = \mathbf{88.64\%}$. Therefore, Amy's level of subject matter knowledge of measures of central tendency is high.

Case Study Five: Tina

Tina was 22 years 9 months old when she was interviewed. At the point of data collection, she was pursuing a 4-year Bachelor of Science with Education (B.Sc. Ed.) program at a public university. She majored and minored in physics and mathematics respectively. Tina obtained an A+ in Mathematics and an A- in Additional Mathematics in her SPM examination (equivalent to O level examination). She also scored an A- in Mathematics in her Matriculation examination.

At the time of data collection, Tina was in the final semester of fourth year studies. She attained 2.93 in the Cumulative Grade Point Average (CGPA) for her first three years of studies in the public university. She does not have any teaching experience prior to this interview. The following sections discuss the findings of Tina's subject matter knowledge of measures of central tendency in each of the four different constructs, namely with reference to context, in handling bias, in problem solving, and in making inference that were emerged from the clinical interview.

Types of subject matter knowledge.

Subject matter knowledge of measures of central tendency with reference to context.

Task 1. Tina's knowledge of the mean is the average. She knew that a decimal average in the given context is illogic but she said that the decimal average can be produced from calculation without any consideration given to the context.

Table 4.107

Excerpt T1

R: How do you think the average was obtained?
S: Total children in all families and then divide by the total number of families. And then you get the average.
R: Can you show and explain using an example?
S: Example if we got family A with 2 children, family B with 3 children. So, the total children is 5 and then divide by the total families which is 2 and you get the average.
R: How can the average be 2.3, and not a counting number like 1, 2, 3, or 4?

S: Families might have even or odd number of children in which we add can give even or odd total and then divide with the total number of families which is also even or odd numbered. So that is why you get 2 point something, 3 point something and not exactly a whole number.
R: Can you explain?
S: This is involving the calculation of average but in reality, the decimal value for children does not exist. I just do not know how to explain this. Ahhh when involving the calculation of average, a decimal value is possible but if we see the logic, decimal value not logic.
R: What does average relate to in relation to the measures of central tendency?
S: Mean
R: Can it relate to another measure besides mean?
S: Let me think. So, far average is mean only. Mode is referring to highest frequency and median is the middle value. So, average is the mean.

Tina explained that the average is obtained when the total number of children in the families is divided with the total number of families. She further explained this in an example saying that if there are two families, family A with 2 children and family B with 3 children, so the total number of children which is 5 is divided with the total number of families which is 2 in order to get the average.

Tina mentioned that a family can have odd or even numbered children. She explained that when these number of children are added for all the families; the total can also be even or odd total. Similarly, she said that the total number of families can be even or odd numbered too. Therefore, Tina mentioned that the division between these even or odd numbered totals can produce a decimal number like 2 point something, 3 point something and not a whole number.

Tina explained that this involved the calculation of an average. She said that in reality a decimal average to represent children should not exist. However, Tina showed difficulty in explaining her point here. She just emphasised that when the calculation of an average is involved, a decimal value is possible but if the logic is considered, the decimal average might not be logic. According to Tina the average relates to only the mean in terms of measures of central tendency. For her, the mode is referring to the highest frequency whereas the median is the middle value. The above Excerpt T1 describes Tina's points on this.

Task 2. Tina utilised the idea of mode in real life situation without the task explicitly stating to do so. She did not utilise the mode as a form of data representation.

Table 4.108

Excerpt T2

R: How would you place the order for the shoe sizes in your shop?
S: I would arrange from the smallest shoe size to the biggest shoe size available. Then I would count the number of shoes sold for each shoe size. For size 4, how many shoes sold and so on till size 9. From here I will determine the shoe size that is sold the most and the shoe size that is sold the least.
.
.
.
R: Can you please explain?
S: For size 4 there is 5, size 5 is 12, size 6 is 5, size 7 is 7, size 8 is 2 and size 9 is 2 also. I order more for sizes that has more demands
R: If you want to choose one female shoe size, which shoe size would you choose?
S: Size 5
R: Why?
S: Because of the highest frequency.
R: Can the shoe size represent the female shoe size in your shop?
S: Cannot
R: Why?
S: Even though size 5 has the highest frequency but the other shoe sizes also have demand. Size 5 does not represent the overall.
R: What would you use as the average shoe size?
S: Size 5
R: Why size 5?
S: Because according to the given data size 5 is purchased the most, so if average is considered than size 5 should be the average.
R: What is the average shoe size?
S: If calculate it is 5.8.
R: How did you get 5.8?
S: Size 4 times with 5, size 5 times 12 and do the same for the rest. {plus all} and divide by the total frequency so you get 5.8.
R: Earlier you mentioned that the average shoe size used is 5. Why?
S: Yes, because just now I thought the average is for the shoe size that ladies wore the most. But if the average is for the shop then it is 5.8.
R: Can you explain?
S: If were to consider the actual average than it is through the calculation which is 5.8.
R: Why?
S: Because when you total up all and divide by the total shoe size will give the proper average.
R: Why not 5?
S: 5 is actually not representing all but only the most bought shoe size and not considering the overall. I thought earlier the average is for the most worn shoe size. So, it is 5 because 5 is worn the most.

Tina mentioned that in order for her to place the order for the shoe sizes in her shop, first she would arrange the shoe sizes from the smallest to the biggest size. Then she would count the number of shoes sold for each shoe size. She explained that for shoe size 4, how many sold and so till shoe size 9. From here, she said that she would determine the shoe size which was sold the most and the shoe size which was sold the least.

Tina explained this further in detail. She said that that size 4 had 5 purchases, size 5 had 12, size 6 had 5, size 7 had 7, size 8 had 2 and size 9 also had 2. Therefore, she said that she would order more on sizes that had more demand.

Tina mentioned that if she had to choose one female shoe size, she would choose size 5 because size 5 had the highest frequency. However, Tina said that she would not use size 5 to represent the female shoe size in her shop. According to her, although size 5 has the highest frequency but there is also demand for the other shoe sizes. Therefore, she said that size 5 does not represent the overall shoe sizes in her shop.

On the other hand, Tina mentioned that she would use size 5 as the average shoe size because size 5 is purchased the most. However, when Tina was asked on what is the average, she mentioned that from the calculation is 5.8. Tina elaborated on the mean calculation when she was asked to explain on how she obtained 5.8. She explained that 4 is multiplied with 5, 5 is multiplied with 12, and so. All the products of the multiplications are added and then divided with the total frequency to get the average as 5.8.

Tina clarified that the reason she mentioned that she would use 5 as the average earlier was because she thought that the average was for the shoe size that ladies wore the most. However, she said that if the average is for the shop then it would be 5.8. Tina clarified that the actual average is through calculation which is 5.8. She explained that the proper average should be obtained by "total up all and divide by the total shoe size". However, she further explained that 5 is not representing the overall but only the most bought shoe size. She made it clear that she thought the average earlier was for the most worn shoe size and 5 is worn the most which is why she mentioned 5 because 5 is worn the most. The above excerpt T2 explains Tina's points on this.

Task 3. Tina identified the mode based on the highest number of comments. Initially, she mentioned that the mode can represent the number of comments made by the students on that particular day but later she changed her mind after realizing that the mode was based on one particular student's comments. She mentioned that the mean is more suitable because it represents the overall number of comments made by all students.

Table 4.109

Excerpt T3

R: What is the mode?
S: 22
R: How did you obtain this value as the mode?
S: The highest value from the table is 22. So, this is the mode.
R: If you were to represent one value for the number of comments made by the students on that day, would you use this value?
S: Yes, can.
R: Why?
S: Because the value represents the most comments made so it can represent the number of comments. The most can be like a voice to the other students. Because I relate with logic, normally comments let say after a talk there is normally feedback is given. From the feedback, we evaluate how the speaker is whether ok or not. So, we look at the comments if more say ok then we can say that the speaker was good and vice versa. [Looked uncertain] One student only. It is based on one student only. In that case, cannot.
R: Why?
S: Because it is based on one student. So, you cannot also represent the number of comments based on this value also. Must find the average.
R: How?
S: Add all and then divide by eight
R: Can you show me?
S: (Writes down the answer in the task sheet) 4.87
R: What if the number of comments made by student D is no longer 22 but 6, what is the mode now?
S: 6
R: How did you find the mode?
S: Based on the highest number of comments
R: If you were to represent one value for the number of comments made by students on that day, would you use this value?
S: I have to count again (she calculates mean by total the number of comments and divide by 8) 2.6. No. I would not use 6 but I would use 3.
R: Why not 6?
S: Because that value represents only one student. By right it should consider all students. That is why we need to consider the mean because it represents all.

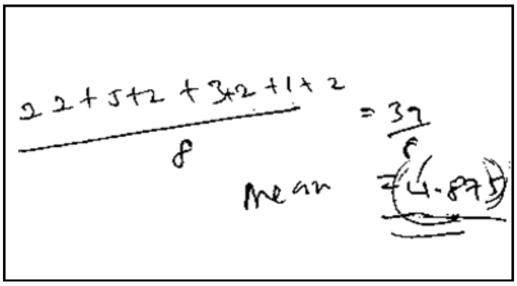


Figure 4.25. Entry T1

Tina mentioned that the mode is 22 in which she based on the highest value given in table. Initially, she mentioned that the mode can be used to represent the number of comments made by the students on that day. According to her, the mode represents the most comments which is like a voice for the other students. She related this with a scenario on the feedback received after a talk. She said that by looking at the comments, if more say ok then we can say that the speaker was good.

However, Tina looked uncertain when she realised that the number of comments for the mode was based on only one student. Therefore, later she changed her answer as the mode cannot represent the number of comments made by the students on that day. She explained that it is based on one student therefore it cannot represent the number of comments made by the students. Instead she said that to find the average. She explained that the average can be obtained by adding all the number of comments and then divide by eight. She got the average as 4.87 which she wrote the calculation for the average in the task sheet. This is shown in Entry T1.

The task was later modified where the number of comments for student D was changed from 22 to 6. Tina mentioned that the mode now is 6 which she still based on the highest number of comments. Tina said the she needed to count when she was asked if she would use this value to represent the number of comments made by students on that day. She performed the mean calculation and got the answer as 2.6.

After obtaining the mean, she mentioned that she would not use the mode to represent the number of comments made by the students on that day.

Instead she said that she would use 3. She mentioned that 6 which is the mode represents only one student. She elaborated that in order to represent something, all the students should be considered. Therefore, she said that the mean is more appropriate because it represents all. The above Excerpt T3 illustrates Tina's points on this.

Task 4. Tina's knowledge of the average is the mean. However, Tina mentioned that the average can be based on the mode but not for this situation that had to deal with the information on a group of students who watched more than 12 hours of TV per day.

Table 4.110

Excerpt T4

R: What does "average" mean in this sentence?
S: Average means if one student how many hours of TV does the student watch. One student represents how many hours.
R: How do you think they got this average 3 hours of TV per day?
S: So, I think the average was obtained by totalling the number of hours students watched TV per day and then divide with how many students involved and you can obtain the average as 3.
R: Can you explain using an example?
S: Student A 2 hours, student B 3 hours, student C 4 hours, total the hours and then divide by the total number of students.
R: How about the information on a small group of students watched 12 hours of TV per day?
S: It would have been totalled up with all the other values and then divide with the total number of informants.
R: How?
S: Say for example there are two or three students who watched 12 hours of TV per day so we take the 12 hours for each one and then add with the other values taken from other students. We total up all and then divide by total number of students.
R: Do you think there is other way they would have got 3 hours of TV per day?
S: Ok, [thinks for a while] {looked unsure} Maybe observe a student watching TV per day for a week and then note the most frequent number of hours watched. Say Sunday the student watches for 1 hour, Monday 2 hours, Tuesday 2 hours, Wednesday 2 hours, 2 hours on Thursday and the rest of the days is 3 hours. So, we take 2 hours as the estimated average.
R: Can you explain further?
S: Say for Sunday 1 hour, Monday 2 hours, Tuesday 2 hours, Wednesday 2 hours, 2 hours on Thursday and the rest of the days is 3 hours. So, 2 hours is most frequent so we can assume the average as 2 hours.
R: What does this value relate to?
S: It depends on the highest frequency which is mode. If this the median and this is the mode.
R: You said the earlier one as median?
S: Oh sorry. The mean and this one is the mode.
R: Can average be based on the mode?

S: I think can.
R: How about the information on 12 hours?
S: If using the mean, you total this information with the rest and divide with the number of students.
R: How about the mode?
S: I am not sure. I do not know how they deal with the information on a group of students watched more than 12 hours of TV per day. I think here in this situation we cannot use the mode.
R: Why?
S: Because the mode does not consider all. Mean considers all. So, cannot be based on the mode.
R: Do you think there are any other ways to obtain the average beside the ones that you explained just now?
S: [Thinks for a while] I think that is all. Only mean.

Tina mentioned that the average means the number of hours of TV that one student watched. She explained that the average 3 hours was obtained by adding the number of hours students watched TV per day and then divided the total with the number of students involved. Tina elaborated this further in her example where she mentioned to total the number of hours for student A who watched 2 hours, student B who watched 3 hours, student C who watched 4 hours and then divide with the total number of students.

She mentioned that the information on a group of students who watched 12 hours of TV per day would have been included in the calculation too. She explained this in an example saying that if there were two or three students who watched 12 hours of TV per day, these 12 hours from each of the students are added with the other number of hours from the other students and then divided with the total number of students.

However, Tina looked rather unsure when she was asked if there were other ways to obtain the average. She thought about it for a while. After when she explained that maybe observe a student watch TV for a week and take note the student's most frequent number of hours watched. She elaborated this further in an example as the following if the student watched 1 hour on Sunday, 2 hours on Monday, 2 hours on Tuesday, 2 hours on Wednesday, 2 hours on Thursday and the rest of the days is 3 hours. Therefore, the average is 2 hours which is based on the most frequent value. Tina mentioned that this related to the highest frequency which is the mode.

Tina mentioned that average can be based on the mean and the mode. She explained that if the mean is considered, the information on a group of students who watched more than 12 hours of TV per day will added with the rest of the information on the number of hours and then divided with total number of students. However, Tina was not sure on how to deal with this information if the mode is considered. She quickly changed her mind that for this situation, the mode cannot be considered because its lack in considering all the data as compared to the mean that considers all the data. Tina was not aware of any other ways of considering the average except the mean. The above Excerpt T4 describes Tina's points on this.

Task 5. Tina mentioned both the mean and the mode can be used to represent the weight of all ten children. However, Tina claimed that the median cannot be used because of an extreme weight in the data.

Table 4.111

Excerpt T5

R: Which is the weight for the median child?
S: 19
R: How did you get 19?
S: Arrange from the smallest number to the biggest number and then take the middle number.
R: Is there any other way to get the median?
S: No only this way. But then for this question the total is odd so we take the middle number straight. If even number, take the two middle numbers add both the numbers and then divide by two. You will get the median.
R: Which is the median if we include another child who weighs 43kg?
S: 19 also.
R: How did you get 19?
S: Because both the middle numbers are 19 and 19. So we add both and divide by two is still 19.
.
.
.
R: But the question is actually asking if it is adequate to use the mean?
S: Ahhh is it? Oh. Yes. To represent the weight of all the 10 children. Yes, we can use mean because mean can be used to represent the weight of all the 10 children.
R: Why?
S: Because mean represent all. If we take median it is the middle value only and not all have the median weight. 43 is far from 19 so we cannot take median. That is why we need to find the average so that the value will not be far away.
R: Is there any other way that we can use to represent the weight of all 10 children?

S: We can use mode. We consider the most weight.
R: What is the mode?
S: 19
R: Why?
S: Because the most can be used to represent.
R: Any others?
S: No only this two only. Mode and mean.

Tina mentioned that the weight of the median child is 19 kg. She explained that she got this value after she arranged the given values from the smallest to the biggest and then she took the middle number as the weight of the median child. When she was asked if there were other ways to get the median, she mentioned that there are no other ways and the one that she explained is the only way. However, Tina added that if the total number of data is odd then the median would be the middle number and if the total is even numbered than the median would be the two middle numbers added and then divided by two.

Tina mentioned that the median after 43 kg is included is still 19 in which she explained was obtained after the two middle numbers in this case both 19 were added and divided by two. Tina said that the mean can be used to represent the weight of all ten children because the mean represents all. She elaborated that the median is not suitable because it is just the middle value and not all have the median weight. She added that since 43 kg is far from 19 kg, therefore the median cannot be considered. She added that the average need to be calculated, so that 43 kg would not be far from the average.

However, Tina mentioned that the mode can be used to represent the weight of all ten children because the mode considers the most weight. She explained that the mode which is 19 is the most and can be used to represent the data. Tina said that the mode and the mean are the only two ways to represent the data in this task. The above Excerpt T5 illustrates Tina's point on this.

Task 6 and task 7. Tina utilised the backward mean calculation to place the prices for all the bags of crisps. Tina conveniently utilised the average price for all the seven bags of crisps where there was no restriction on its use. However, when there was a restriction posed on the use of the average price again, Tina randomly picked the first price in which she subtracted the value from the total price value. She repeated this for all the other prices too until the last price which gave her a zero balance.

Table 4.112

Excerpt T6

R: How did you place the price for each of the bags?
S: I placed all the prices as RM 27.
R: Why?
S: First I found the total. The average price is 27. Here there are 7 places. So, 27 times by 7, so the total is 189. Therefore, when you place all 27 and by adding all 27, you can get the total 189.
R: How did you base each one of this value?
S: I just take the value which is given. But if you want to place other values also can. As long as the total for all prices is 189 and then after the total is divided by 7, you get 27.
R: Why 27?
S: Because that is the given value and when we multiply with 7 we get 189 and 189 divide by 7 again, we get 27 so place 27 in each of the boxes. The easiest way.

Tina placed the prices for all the bags of crisps as RM 27. She elaborated that first she found the total for all the prices. She multiplied 27 with 7 and got the total as 189. She explained that if she placed all the prices as 27 and then she added these prices, she would get 189. Tina mentioned that she took the given average value as the value to be considered for all seven bags of crisps. However, she said that other prices can be considered too as long as the total for all the prices is 189 and when the total is divided with 7, gives the average 27. Tina also mentioned that she placed 27 as the price for all seven bags of crisps because it was the easiest way. This is shown in the above Excerpt T6.

Table 4.113

Excerpt T7

R: Look at Task 7. How did you place these values?
S: It can be any value. I randomly placed these values. When the values are added up, it gives 189. So, 189 divide by 7 to get 27

R: How did you base each of the values here? S: I take randomly. From 189 minus the first value 29, then the rest one by one till the last value that gives a zero balance from the total. As long as when we add up these values is 189. Just make sure that none of the bags is 27. We still get 27 as the average price. R: Any specific way to decide on the values? S: Mmm…any how the question did not specify anything so that is why I placed it randomly. But to make it easier, I just find first a number that ends with 9 so that I minus 189 with this random number that ends with 9 then the rest of the values I estimate roughly. This is to make my estimation of the other values easier.

In Task 7, Tina placed the prices randomly. She utilised the backward mean calculation to obtain the total price for all seven bags as 189 in which she said when 189 is divided with 7, 27 is obtained. Tina explained that she placed the prices randomly but she utilised the total 189 to minus the first random value that she took which was 29. She repeated this with the rest of the values till the last value for the seventh bag that gives a zero balance from the total. She said that she just made sure that all the prices add up to 189 and none of the bags is 27. She mentioned that this would give the average as 27.

Tina mentioned that since the question did not specify anything, therefore she placed the prices randomly. However, to make things easy for her, she found the first number that ended with 9. She explained that if she subtracted the first random number that ended with 9 from the total 189, the rest of the values she could estimate easily. Excerpt T7 describes Tina's point on this.

Summary.

Table 4.114

Subject matter knowledge with reference to context

Task	Statistical element	Conclusion
Task 1	Mean as average	Tina's knowledge of the mean is the average. She elaborated that the average is obtained when total number of children is divided with the total number of families. Tina also explained on the mean calculation in her example. Tina said that any division involving two even odd numbers in the calculation of average can produce a decimal number. According to her, a decimal average is possible but might be logic. Tina related average to the mean and said that it cannot be referred to any other measures of central tendency.

Task 2	Idea of mode	Tina utilised the idea of mode without the task explicitly stating to do so. She used the idea of frequency tally and determined the mode based on the most purchase.
	Mode as data representation	Tina said that the mode cannot be used to represent the data because the mode size although showed the highest demand but she said that there are also other shoe sizes that have demand too. She said that mode size does not represent the overall.
	Mode as average	Although Tina initially used the mode as the average but later she explained that the mode cannot be used as the average. She explained that she thought the average was for the most purchased shoe size. However, she changed her mind after knowing the average was for the shop. Tina mentioned that the average had to be calculated using the mean calculation. She explained that the mode cannot be utilised as the average because it is not representing the overall and only representing the most.
Task 3	Idea of mode	Tina obtained the mode as the highest value shown in table. Her knowledge of the mode is numerical.
	Mode as data representation	Tina said that the mode cannot be used to represent the data because it is based on one particular student. Instead she said that the average can be used to represent the data. She explained that the average can be obtained based on the mean calculation. Tina also mentioned that the mean is more suitable because it represents all students.
Task 4	Median as average	Tina mentioned that the average was obtained from the mean calculation. She elaborated this further in her example too. Regarding the information on the extreme data, she said that the information would be included in the calculation too. However, Tina also said that the estimated average can be obtained based on the idea of most or mode. However, she changed her mind after she was probed on how the extreme data would have been handled if the mode was considered. She realized that the mode does not consider all the data like the mean does. Therefore, she now claimed that the mode cannot be used. She said that only the mean can be used because it considers all the data.
Task 5	Idea of median	Tina carried out the median procedure by arranging the data in ascending order first. She took the middle number as the median for the case where the data is odd numbered. For the case where the data is even numbered, she said that the median is obtained after the two middle numbers are added and then divided by two.
	Robustness of measures	Tina said that the mean is adequate to represent the set of data that contained an extreme value. She explained that the mean represents all the data even the extreme. She said that such mean value would not be far from the extreme because the value calculated included the extreme. She said unlike the median which only considers the middle value. However, Tina said that the mode can be used to represent the weight of all ten

		children. She mentioned that the idea of the most can be used to represent.
Task 6	Idea of average	Tina placed the values of all seven bags of crisps as 27. Tina utilised the backward mean calculation to find the total price. She said that when all the prices of the bags of crisps are placed as 27, these prices when added back will give the total price. She explained that 27 times 7 gave her 189 as the total price and when 189 is divided with 7 again, she got 27 in which she placed this value for all seven bags. Tina said this was the easiest construction.
Task 7	Idea of average	Tina placed the prices randomly. She said that she made sure that when all the seven prices were added, she got the total price 189. She said that she took the first random value as 29 which she subtracted from 189. She wanted the first random value to end with 9. The rest she said she could estimate roughly based on the balance total. She said that she also subtracted one by one the rest of the values from the balance total until she got the balance zero after the seventh value was subtracted.

Subject matter knowledge of measures of central tendency in handling bias.

Task 8 and task 9. Tina obtained the mode based on the highest frequency, the median was the middle data after all the data were rearranged in ascending order and the mean calculated from the total of all data divided with the total number of data. Tina mentioned that the mean is the best measure to represent the data. Although Tina had calculated the mean based on all given data but she referred the best measure mean as the mean calculated after excluding the error.

Table 4.115

Excerpt T8

R: What are the measures of central tendency that you can obtain based on the above data?
S: The mode is 6.3.
R: How did you get the mode?
S: 6.3 occurs three times. 6.3 has the highest frequency. The others 6.0 occurs twice, 6.1 occurs once, 6.2 occurs once, 15.3 occurs once and 6.15 occurs once. So, 6.3 occurs thrice so that one is the mode. For the median, arrange all the numbers and take the middle number which is 6.2.
R: How would you arrange the numbers?
S: Arrange from the smallest number to the biggest number.
R: Any other way that you can find the median?
S: For this case only this way
R: How about the mean?
S: The mean. I total up all the numbers and then divide by nine because there are nine numbers. I get 7.18
R: Which is the best measure of central tendency to represent the actual weight of this object?
S: For me is the mean which is 7.18
R: Why?

S: Because relate to the experiment, what I usually do is that I repeat experiment for three times and then find the average because sometimes there might be an error. So, from there we can get the actual weight. Same goes to this one. If we total all the values and try to get the average value because there might be an error in weighing the object.
R: You mentioned that if there exist an error. Can you explain?
S: There might be an error here because the weight is taken on a weighing scale. So, error is possible.
R: How about this data?
S: The values in the data are close together, except for one value which is 15.3. Only one value that is far. By right this value, should not be included in the calculation. If it is not included, then the mean will be different. So, we add the rest excluding this value and then divide by eight. This mean will be close to the other values in the data.
R: How then regarding the mean?
S: For mean, I would not consider this value. So, I will take the eight values. And then for the mean, add up the eight values and divide by eight. I get 6.17.
R: Which mean would you consider now?
S: I would consider 6.17
R: Why do you think the mean is the best to represent the data?
S: Because the mean considers all of the values. If I take the mode, I only consider the value that carries the highest frequency only. If the median, it is the middle value only.
R: Would you consider the mode or the median to represent the actual weight?
S: I don't think so.
R: Why?
S: Because as I mentioned earlier. The most acceptable is the mean. If mode, it refers to the most frequent and not the others so cannot also. If median, the middle value. So, we should not consider the mode or the median. The mean 6.17 is the most suitable.

Tina mentioned that all three measures of central tendency can be obtained based on the given data in this task. She found the mode as 6.3 which she based on the highest frequency because 6.3 occurred thrice more than all the other data that occurred at most twice. Tina said that the median is 6.2 in which she found by first arranging all the data from the smallest number to the biggest number. Then she took the middle number which is 6.2 as the median. She said that this was the only way to find the median.

For the mean, Tina mentioned that she totalled up all the numbers and then divided the total with nine. She got the mean as 7.18. Tina mentioned that the best measure of central tendency to represent the actual weight of this object is the mean. She explained that if she related this to an experiment, normally the experiment is repeated a few times and the readings are recorded. Based on these readings the average is calculated. According to Tina, this takes into consideration of any error in the readings and the likelihood of getting the most reliable result. Similarly, to the weights. She mentioned that all the weights are averaged because there might be an error in weighing the weight of the object.

When Tina was probed to explain further on her response that touched on error, she elaborated that in weighing the weight of an object an error is possible. She said that in the given data, all the data values are close to one another except for one, 15.3 which was far from the rest. According to Tina, 15.3 should not be included in the calculation of the average. She added that if 15.3 is excluded than the mean would be different. The mean now would be calculated based on eight data and this value now which she calculated as 6.17 would be closer to the remaining eight data.

Tina mentioned that she would consider 6.17 to represent the actual weight of the object. She explained the reason why she considered 6.17 because the mean takes into account all the data unlike the mode that considers only the data with the highest frequency whereas the median considers only the middle value. Thus, this was the reason that Tina did not want to consider the mode or the median to represent the actual weight of the object. Excerpt T8 describes Tina's points on this.

Table 4.116

Excerpt T9

R: What are the measures of central tendency that you can obtain based on the above data?
S: Median and mean. Median 54 000 just arrange from the smallest number to the biggest number. 42 000, 48 000, 54 000, 60 000, 180 000. So, the middle number is 54 000 is the median. Mean, total up all the five numbers and divide by five so you get 76 800.
R: What about the mode?
S: All five employees got different salaries. So, no mode.
R: Which is the best measure of central tendency to represent the actual salary of the company?
S: Mean.
R: Why?
S: Same reason it considers all.
R: Can you explain further on the mean?
S: The data is similar to the one before. Employee E has the highest salary. Too much difference compared to the other employees. If the E is not included and we find another mean, it can represent. The mean is 51 000
R: How?
S: We don't take count of the E. Total up for the other four values and then divide by four. Exclude E
R: Why?

S: Because E has quite a big difference as compared to the other values. If we add this, it cannot represent the salary of the company. If we see the salary is in the range of 40 000 to 60 000 but then E is 180 000. There is quite a bit of difference. So, we cannot take count of this.
R: Which mean will be used to represent the actual salary?
S: 51 000
R: Why?
S: Because it is around the other salaries excluding 180 000
R: How about median?
S: No. Median is the middle

Tina mentioned that for the given data in this task, the median and the mean can be obtained. She found the median as 54 000. She said that she arranged the data from the smallest to the biggest: 42 000, 48 000, 54 000, 60 000, 180 000 and found the median from the middle number in the arrangement. Tina calculated the mean as 76 800. She said that she totalled up all five salaries and then divided the total with the number of employees, five. Tina mentioned that the data in this task has no mode because all the employees had different salaries.

According to Tina, the best measure to represent the actual salary is the mean. She explained that this is because the mean considers all the data. However, when Tina was probed to explain further, she elaborated that employee E has the highest salary which was far different from the other salaries. She said that if E is not included and the mean is found based on the remaining four salaries then the new mean which is 51 000 can represent the actual salary.

Tina said that Employee E's salary 180 000 has quite a big difference compared to the other salaries. If E's salary is added into the mean calculation, the mean cannot represent the salary of the company. She said that the other salaries are within the range of 40 000 to 60 000. Thus, Tina said that the salary of employee E should not be included. She said that the mean based on the remaining four salaries which is 51 000 represents the actual salary because it is within the range of the other salaries. Tina also mentioned that the median cannot be utilised to represent the actual salary because it only refers to the middle value. Excerpt T9 describes Tina's points on this.

Summary.

Table 4.117

Subject matter knowledge in handling bias

Task	Statistical element	Conclusion
Task 8	Identifies and summarises data using mean	Tina found the mean by adding all the nine data and then divided the total with nine.
	Identifies and summarises data using median	Tina identified the median by first arranging all the data in ascending order and then she located the middle value as the median.
	Identifies and summarises data using mode	Tina identified the mode based on the highest frequency. The data point that occurred the most is considered as the one with the highest frequency.
	Best data representation measure	Tina mentioned that the mean is the best measure to represent the data. However, Tina said that the mean calculated after excluding the extreme value in which she believed to be an error would be the best mean to represent the data. She explained that the mean calculated after excluding the extreme is close to the rest of the data unlike the earlier one. Tina said that she would not consider the mode or the median to represent the data because the mode is based on the value with the highest frequency meanwhile the median is just the middle value. According to her, the mean is the most suitable measure because it considers all the data.
Task 9	Identifies and summarises data using mean	Tina found the mean by adding all the five data and then divided the total with five.
	Identifies and summarises data using median	Tina identified the median by first arranging all the data in ascending order and then she located the middle value as the median.
	Identifies and summarises data using mode	Tina said there is no mode for this situation because all the data occurs only once so there is no data with the highest frequency.
	Best data representation measure	Tina mentioned that the mean is the best measure to represent the data because it considers all the data. However, she said that the mean calculated after excluding the extreme data would be the most suitable mean to represent the data because the mean value is around the rest of the data.

Subject matter knowledge of measures of central tendency in problem solving.

Task 10. Tina obtained the total grade point by multiplying the given GPA with the number of semesters involved for both the colleges. After when she added the total grade points of both colleges to obtain the overall total grade points in which later she divided this with five semesters to get the overall GPA for the students.

Table 4.118

Excerpt T10

R: How did you find the GPA?
S: For A there are two semesters involved with GPA 3.2, the total is 3.2 times 2 which is 6.4. Same goes to B, there are three semesters so 3.8 times 3 which is 11.4. To find the overall GPA, for five semesters, we need to add the total for two semesters which is 6.4 and 11.4 and then divide by 5.
R: Why did you times 3.2 with 2?
S: Because the given GPA is for two semesters and I want to find the total for A.
R: Why did you times 3.8 with 3?
S: Same like just now. The GPA is for three semesters and I want to find the total for B.
R: Why did you add 6.4 and 11.4?
S: Because I want to find the overall GPA, so I add the total for A and B and I find the overall total.
R: Why divide by 5?
S: Because of five semesters. So, I divide the overall total with five and I get the overall GPA as 3.56.

Tina mentioned that for college A there are two semesters involved and the given GPA is 3.2. Therefore, she said that the total grade points for college A is 6.4 in which she calculated by multiplying 3.2 with 2. She said that the same applied to college B. There are three semesters involved with the given GPA as 3.8. Thus, Tina said that the total grade points for college B is 3.8 times 3 which is11.4. Tina explained that to find the overall GPA for five semesters, both the total grade points, 6.4 and 11.4 are added and then divided by five because there is a total of five semesters involved. From here Tina obtained the overall GPA as 3.56. The above Excerpt T10 describes this.

Task 11. Tina found the total hours of TV watched for both rural and urban groups from the given averages. She added the group total to get the overall total in which later she divided with 100 students. This gave Tina the average TV viewing time per weekend for all 100 students.

Table 4.119

Excerpt T11

R: How did you get this average?
S: First find the total hours of TV viewing by the students. For the rural 25 times by 8 so the total hours is 200 hours. For the urban is 75 times by 4 which is 300 hours. So, to find the average TV viewing time for all 100 students, add 200 with 300 divide by 100 students. We will get 5 hours TV viewing time per weekend.
R: Why are you dividing by 100?
S: Because of 100 students
R: Why 25 times 8?
S: To find the total TV viewing time for rural students.
R: Why 75 times 4?
S: To find the total TV viewing time for urban students.

Tina explained that she had to find first the total hours of TV viewing for the students in both groups. For the rural students, she said that she multiplied 25 with 8 to get the total of TV viewing time as 200 hours. Similarly, she multiplied 75 with 4 to get the total TV viewing time for urban students as 300 hours. Tina mentioned that to find the average TV viewing time for all 100 students, she added the group totals. She added 200 with 300 and then divided the total with 100. Tina obtained the average as 5 hours of TV viewing time per weekend. Tina explained that the total hours was divided with 100 because the average was for all 100 students. The above Excerpt T11 illustrates Tina's points on this.

Summary.

Table 4.120

Subject matter knowledge in problem solving

Task	Statistical element	Conclusion
Task 10	Backward mean calculation	Tina solved for the weighted mean for Task 10 without the task explicitly stating to do so. She had carried out the backward mean calculation, where she obtained the total for each group of data set from the given group means.
	Representative nature of the mean	Tina used the knowledge related to the representative nature of the mean in making sense of why data sets need to be combined in the weighted mean.
	Forward mean calculation	Tina also carried out the "forward" mean calculation to calculate the weighted mean.
Task 11	Backward mean calculation	Tina solved for the weighted mean for Task 11 without the task explicitly stating to do so. She had carried out the "backward" mean calculation, where she obtained the total for each group of data set from the given group means.
	Representative nature of the mean	Tina used the knowledge related to the representative nature of the mean in making sense of why data sets need to be combined in the weighted mean.
	Forward mean calculation	Tina also carried out the "forward" mean calculation to calculate the weighted mean.

Subject matter knowledge of measures of central tendency in making inference.

Task 12. Tina had several ways to select the player. She said that the selection can be based on the lowest and the highest score comparison. She also mentioned that the average score can also be compared to decide on the best player. However, according to Tina among all these ways, the average score comparison is still the best method to be used in the selection of the player.

Table 4.121

Excerpt T12

R: If Coach Ahmad can only select one of the two players, who would you recommend he select?
S: Beng.
R: Why?
S: Because of various reason. First one. We look at the scoring results, between Beng and Ramu. Beng scored the highest which is 28 meanwhile Ramu the highest is only 23. So we can see Beng did better. If we look at the lowest score also, Ramu got 16 meanwhile Beng got 18. So Beng did better than Ramu. If we calculate the average score, Beng is better than Ramu. So Beng is better to be chosen.
.
.
.
R: You mentioned several ways. Which is the best way to use to select the player?
S: Mmm… [thinks] average.
R: Why average?
S: Average will consider all the scores because not all the time the highest is scored by the best player. It happens to be that in this situation Ramu's highest score is less than Beng's and so was his lowest score was less than Beng's. Average will still be better and here Beng's average is higher than Ramu's average. So choose Beng.

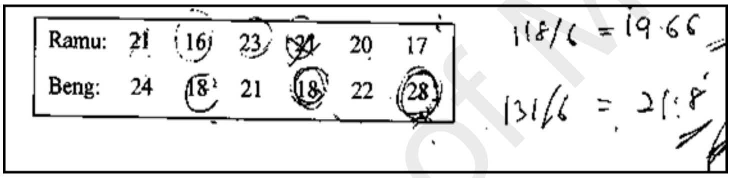


Figure 4.26. Entry T2

Tina mentioned that she would recommend Beng to be selected. She said that there are various reasons to select Beng. One of the reasons according to her is the players' scores. Tina said that Beng's highest score was 28 meanwhile Ramu's highest score was 23. Thus, she claimed that Beng did better than Ramu.

Tina did similar comparison on the lowest score. She said that Ramu's lowest score was 16 whereas Beng's lowest score was 18. Since Beng's lowest score was higher than Ramu's lowest score, therefore, again Tina claimed that Beng performed better than Ramu. Lastly, Tina compared the average scores of both players and she found that Beng's average score was better than Ramu's average score which is shown clearly in Entry T2. So again, Tina claimed that Beng performed better than Ramu and should be chosen.

According to Tina among all the ways that she mentioned can be used to select the player, the average score comparison is the best. She explained that the average considers all the score. She said that the highest score comparison is not necessarily accurate because not always the highest score is scored by the best player. She said that in this situation, it happened to be Ramu's highest score was less than Beng's. So was Ramu's lowest score was less than Beng's. She said that comparing the average score is better and since Beng's average was higher than Ramu's average, so Beng was recommended to be chosen. The above Excerpt T12 describes Tina's points on this.

Task 13. Tina used mean comparison to select the player for unequal sized data. She believed that the mean or the average score is the best way to compare and select the player when the data is unequal.

Table 4.122

Excerpt T13

R: Now Ain, look at task 13 now. Who would you recommend him to select here?
S: Why is it like this? [Points at the data which is not of equal size]
R: Unequal data set. Here there is 6 scores (points at Khan) and here there is 8 scores (points at Ali).
S: Ok (starts working on the task). The same like the earlier question. Find the mean but for the first one, find the mean divided by eight because of eight matches. For Khan is six matches so divide by six.
R: Who would you recommend him to select?
S: Khan. Because his average is higher than Ali.
.
.
.
R: Khan got six scores and Ali got eight scores. Is it ok to compare the averages?
S: It is ok because the average considers all scores for the player.
R: Do you think the average is the best way to select the player?
S: Yes
R: Why?
S: Because if we compare with the highest score it might be the lucky score that the player scored for that particular match. But when we calculate the average means we are considering all scores in all the matches regardless whether the highest or lowest scores.

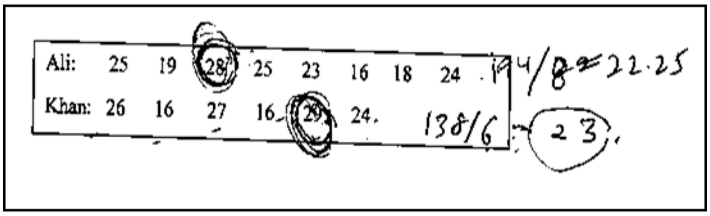


Figure 4.27. Entry T3

Tina showed some confusion over the unequal sized data presented in this task. She thought that the data was insufficient for Khan. After she was told that the data is of such for Khan and Ali, Tina started working on the task. She said that she would find the mean score for both the players. She mentioned that for the first one which was referring to Ali, she would divide by eight because eight matches while for Khan, she said she would divide by six because of six matches.

Tina said that she would select Khan because his average is higher than Ali. Tina's mean calculation for the average is shown in Entry T3. She said that it is ok to compare the averages even if the set of scores for both players are not equal because the average considers all scores for that player. She believed that the average is the best way to select the player. She explained that comparing the scores such as the highest score might not be accurate because it can be a lucky score for that player. Whereas, she said if the average is calculated meaning that all the scores for all the matches regardless whether the highest or the lowest are considered. The above Excerpt T13 illustrates Tina's points on this.

Task 14. Tina compared the mean to choose the class that performed better. She believed that the mean comparison is the best way to determine the class that performed better.

Table 4.123

Excerpt T14

R: Based on the graphs below, which class did better? S: 5A did better. R: Why? S: I take the total scores divide by the number of students. I find the average. So, I got 4.89 for 5A and for 5B is 3.8. And then from the average it shows that 5A did better. . . . R: Besides average anything else that you can use to determine the class that performed better? S: No other ways. For this question, it should be based on the average. R: Why? S: Because it takes count of all. R: What all? S: All the scores for all the students

Tina mentioned that based on the graphs, 5A performed better. She explained that she took the total number of corrects and then divided it with the total number of students. She said that she found the average. According to her, the average for 5A was 4.89 and the average for 5B was 3.8. She compared the average for both the classes in which she said it showed 5A did better. Tina said that for this question, there are no other ways to determine the class that performed better except for the average. She said that the average takes count of all the number of corrects for all students. This is described in the above Excerpt T14.

Task 15. Tina mentioned that before calculating the mean, she actually looked at the graphs and tried to determine the class that performed better. She said that the information from the graphs mislead her decision earlier in which she thought 5C performed better compared to 5D. After when she calculated the mean and she found that the performance of 5D was actually better than 5C.

Table 4.124

Excerpt T15

R: Based on the graphs below, which class did better? S: 5D did better. R: How? S: I did the calculation to find the mean and the mean for 5D is better than 5C. .

.
.
R: Beside mean, is there any other ways to determine which class did better?
S: Before calculating the mean and by looking at the graphs I thought this class (pointing to 5C) did better because the number of the students and the scores are quite high but based on the mean this class (pointed at 5D) is better.
R: Why did you think that based on the graph 5C did better?
S: Because there are lot students who scored 5 and 6, the mode. Compared to this one (pointed at 5D). But this one (5C) there are 6 students who scored 7 corrects same as 5D's mode. So, from here I knew that I cannot refer to the graphs. I need to find the average.
R: How about the size of the data?
S: 5C involves lot students as compared to 5D. Although this one is 21 (5D) and this one is 36 (5C) but dividing with the number of students gives you that this one (5D) did better. This is actually the most appropriate way to determine which class did better. If we look at the graphs it is not so accurate because involve different number of students.

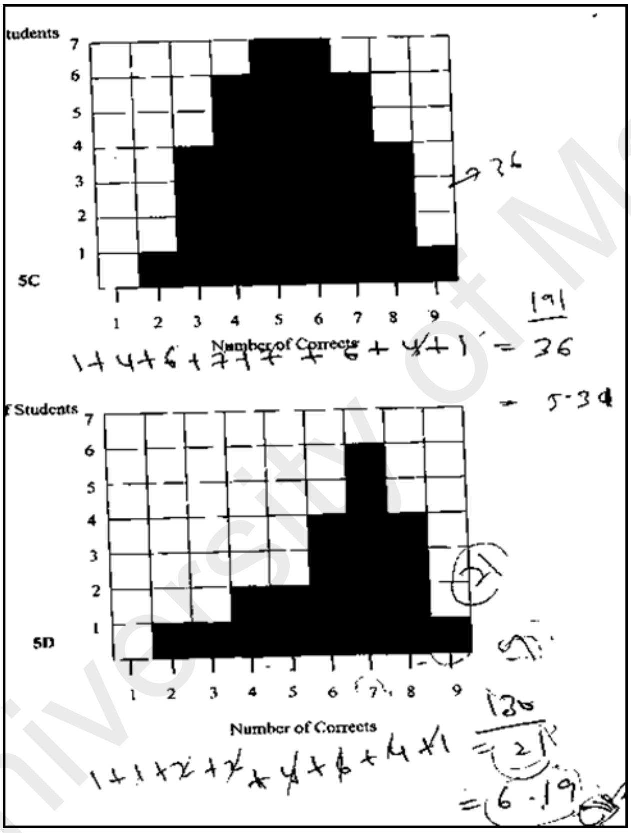


Figure 4.28. Entry T4

Tina mentioned that 5D did better. She said that she calculated the mean and the mean for 5D was better than the mean for 5C. When Tina was probed on other ways to determine the class that performed better, she mentioned that she looked and compared the graphs first before calculating the mean. She said that initially she thought that 5C did better because the number of students and the scores were quite high but based on the mean, she found that 5D was actually better. This is shown in Entry T4.

She explained that she noticed that a lot of students scored 5 and 6, the mode scores as compared to 5D. However, she also noticed for 5C there were 6 students who scored 7 corrects which was actually the mode score for 5D. Therefore, Tina knew she cannot determine the class that performed better by referring to the graphs only and that she needed to calculate the average.

Moreover, she also mentioned that 5C had more students (36) as compared to 5D (21). However, she said that in the calculation for the average, dividing the total number of corrects with the number of students will give the class that performed better. She found that 5D performed better. Tina elaborated that this is the most appropriate way to determine the class that performed better because the data involved unequal number of students. Tina also claimed that the decision about the class performance based on the graphs might not be that accurate. Excerpt T15 describes Tina's points on this.

Summary.

Table 4.125

Subject matter knowledge in making inference

Task	Statistical element	Conclusion
Task 12	Summarises equal sized numerical data using measures of central tendency	Tina summarised equal sized numerical data using the mean.
	Utilises the appropriate measure to compare and draw conclusion about two equal sized numerical data	Tina mentioned that there are several ways to compare and select the player. She said that the lowest and the highest score and the average can be compared. However, Tina said that among the methods that she mentioned, the average is the best way to compare and select the player. According to Tina, the average considers all the scores and not only the highest or the lowest score.

Task 13	Summarises unequal sized numerical data using measures of central tendency	Tina summarised unequal sized numerical data using the mean.
	Utilises the appropriate measure to compare and draw conclusion about two unequal sized numerical data	Tina mentioned that the mean or the average is best way to compare and select the player because the average considers all the scores of the player regardless the highest or the lowest
Task 14	Summarises equal sized graphical data using measures of central tendency	Tina summarised equal sized graphical data using the average or the mean.
	Utilises the appropriate measure to compare and draw conclusion about equal sized graphical data	Tina compared and decided on the class that performed better using the mean comparison. According to her, the mean or the average takes count of all scores for all students in each class.
Task 15	Summarises unequal sized graphical data using measures of central tendency	Tina summarised two unequal sized graphical data using the mean.
	Utilises the appropriate measure to compare and draw conclusion about unequal sized graphical data	Tina mentioned that she tried comparing the two classes based on the graphs. However, she found that the graphs were misleading and later she decided to calculate and compare the average or the mean to decide on the class that performed better. Tina mentioned that the graphs cannot be compared because of the differences in the number of students in both the classes. Instead, she said that the average can deal with this unequal number of students for both the classes and is more appropriate.

Level of subject matter knowledge of measures of central tendency.

Level of subject matter knowledge of measures of central tendency with reference to context. In Task 1, Tina mentioned that the average was found by dividing the total number of children in all families and then divide with the total with the number of families. She also showed the mean calculation in her example. Tina knew that for the context of the given data which was referring to the number of children, the decimal average is not logic. However, she said that the calculation of average can result in the decimal average. Tina's knowledge of the average here is the mean. Therefore, Tina presented the knowledge of the mean as the average (P) with appropriate (A) justification.

In Task 2, Tina had utilised the idea of mode without the task explicitly stating so. She found the frequency for each shoe size and had identified the mode as 5 based on the most sold size or the highest frequency. Therefore, Tina's idea of the mode was found present (P) and correct (A).

Tina said that the mode cannot be used to represent the female shoe size in her shop. She explained that size 5 has the highest frequency but there are other shoe sizes that are on demand too. She said that size 5 does not represent the overall. Therefore, Tina's knowledge of the mode as a form of data representation was found to be not present (NP) and the absence is found to be problematic (I).

Tina utilised the mode as the average initially because she thought the average was for the most worn shoe size. However, her later explanation indicated that for Tina the average is the mean. She calculated the average as 5.8. Tina had lacked in the knowledge that the mode can be used as a quick method in the reporting of an average. Therefore, Tina's knowledge of the mode as the average was not only found to be not present (NP) but also problematic (I).

In Task 3, Tina identified the mode as 22. She took the mode as the highest number of comments. When the task was modified where the number of comments for student D was changed from 22 to 6, Tina identified the mode as 6 in which she still referred to the highest number of comments. Tina had identified the mode incorrectly. First, she had taken the highest number of comments to be the mode in which she believed to be highest frequency. Tina also lacked in the knowledge that the mode can be categorical. The mode is referred to be the category that carried the highest frequency. Therefore, Tina's idea of the mode was marked to be not present (NP) and the lack is considered problematic (I).

Initially, Tina mentioned that the mode can be utilised to represent the number of comments made by the students. She said that the most number of comments can be like a voice. However, when she realized that the mode was based on one student's comment, she now claimed that the mode cannot be used to represent the number of comments made by the students on that day. Instead, she said that the average is better. Therefore, Tina's knowledge of the mode as a form of data representation was found to be not present (NP) and the lack was considered problematic (I).

In Task 4, Tina mentioned that the average referred to how many hours of TV per day a student watched. Her explanations revealed that her knowledge of the average is the mean in which she elaborated on the mean calculation in her example. Tina said that in the calculation of the average, all the data including the extreme data would be included in the calculation.

Tina said that the average can be sometimes based on the mode. However, she later said that the mode cannot be used, when she was asked on how the information about the extreme data would have been dealt here. Tina said that the average should be based on the mean because the mean considers all. Tina did not mention if the median can be considered.

The average based on any data involving human population such as the Malaysian primary school students is normally based on the median because the median is definitely a more robust measure when dealing with data that contained extreme data. Therefore, Tina did not present the knowledge of the median as the average (NP). However, Tina's knowledge of the mean as the average is acceptable because an average can be based on any of the three measures of central tendency but her explanation involving the mean calculation that included the extreme data was found to be inappropriate therefore her justification was considered inappropriate (I). The appropriate justification would have been that the mean considered excluded the extreme data.

In Task 5, Tina found the median by rearranging all the weights from the smallest to the biggest number. She found the median weight for nine children as 19 which was the middle weight in the rearrangement. She also found the median weight for 10 children after 43 kg was included in the data. She got 19 as the median after she added the two middle numbers and then divided the total by two. Tina's knowledge of the idea of the median was found to be present (P) and accurate (A).

Tina said that the mean is adequate to represent the weight of all ten children. She said that the mean represents all whereas the median considers only the middle value. She said that 43 is far from the median 19. On the other hand, the average will not be far from 43. According to Tina also, the mode can be used to represent the data because she claimed that the most can be used to represent. Tina's knowledge related to the robustness of the measures of central tendency was found to be not present (NP) and problematic (I) because she lacked in the knowledge of the median being a more robust measure in the presence on an extreme data. She also lacked in the knowledge that the extreme data should be excluded in the calculation of the mean and that the mean calculated after excluding the extreme data would be more representative of the data.

In Task 6 and Task 7, Tina constructed data sets based on the given average value and the conditions imposed. In Task 6, Tina placed the prices randomly but around the average price. She utilised the backward mean calculation and got the total price as 189. She said that she utilised the average value given for all the seven prices because when 27 is multiplied with 7, the total price 189 can be obtained. Therefore, Tina's knowledge of constructing a data set based on the idea of average was found to be present (P) and appropriate (A).

In Task 7, Tina utilised the backward mean calculation and calculated the total price as 189. She then placed randomly the prices. She utilised the total price and subtracted the first random value 29 from the total 189. She did the same with the rest of the values until the seventh value. Tina said that now the balance from the total price was 189 was zero. She said when all the prices were added back, 189 was the total. Tina also mentioned that she just made sure she did not utilise 27 as one of the random prices that she placed. However, Tina said that to make her calculation and estimation easier, the first random price that she chose ended with 9. Tina's knowledge of constructing a data set based on the idea of average for Task 7 was found to be present (P) and appropriate (A).

Table 4.126

Coding rubrics for determining level of Tina's subject matter knowledge of measures of central tendency with reference to context

Task	Subject Matter Knowledge of Measures of Central Tendency with reference to Context																														
	Mean as average				Mode as average				Median as average				Idea of mode				Idea of median				Mode as data representation				Robustness of measures				Idea of average			
	P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP	
	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I
Task 1	X																															
Task 2								X					X											X								
Task 3																X								X								
Task 4												X																				
Task 5																	X											X				
Task 6																													X			
Task 7																													X			
	1												**1**				**1**												**2**			

Legend: P = Present NP = Not Present A = Appropriate I = Inappropriate

Tina's Percentage of Subject Matter Knowledge of Measures of Central Tendency with Reference to Context

$$= \frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)} \text{ X } 100\% = \frac{\mathbf{5}}{\mathbf{11}} \times \mathbf{100\%} = 45.45\% \text{ (Medium)}$$

Level of subject matter knowledge of measures of central tendency in handling bias. In Task 8, Tina calculated the mean for all nine data. She added up all the nine data and then divided the total with nine. She got the mean as 7.18. Therefore, Tina not only identified that the mean can be obtained but also summarised the data using the mean by calculating the mean value. Therefore, Tina's knowledge of identifying and summarising the data using the mean was found to be present (P) and correct (A).

Tina said that the median 6.2 was the middle number after the data was arranged in the ascending order. Thus, Tina's knowledge of identifying and summarising the data using the median was found to be present (P) and correct (A).

Tina identified the mode as 6.3. She said that 6.3 occurred thrice and had the highest frequency. Therefore, Tina's knowledge of identifying and summarising the data using the mode was found to be present (P) and correct (A).

Tina mentioned that the mean is the best measure to represent the actual weight. She explained that the mean 6.17 which is calculated after the error data 15.3 was excluded from the data is considered the best representation measure. Thus, Tina's knowledge related to the best representation measure is found to be present (P) because the mean can be used as the best representation measure. Tina's justification that the mean calculated after excluding the extreme data is more representative of the data is also considered as appropriate (A).

Similarly, in Task 9, Tina calculated the mean for all five data. She added up all the five data and then divided the total with five. She got the mean as 76 800. Tina not only identified that the mean can be obtained but also summarised the data using the mean by calculating the mean value. Therefore, Tina's knowledge of identifying and summarising the data using the mean was found to be present (P) and correct (A).

Tina obtained the median as 54 000 after she arranged the data in ascending order first. The middle data in the rearrangement was identified as the median. Thus, Tina's knowledge of identifying and summarising the data using the median was found to be present (P) and correct (A).

Tina mentioned that there is no mode in the given data because all the salaries of the employees are different. Thus, Tina's knowledge of identifying and summarising the data using the mode was found to be present (P) and correct (A).

Tina mentioned that the mean is the best measure to represent the actual salary. However, the mean that she was referring to as the best representative measure is the one that is calculated after excluding the extreme data. Tina calculated the mean based on the remaining four salaries and got the mean as 51 000. Thus, Tina's knowledge related to the best representation measure is found to be present (P) and her justification was found to be appropriate too (A).

Table 4.127

Coding rubrics for determining level of Tina's subject matter knowledge of measures of central tendency in handling bias

Task	Subject Matter Knowledge of Measures of Central Tendency in Handling Bias															
	Identifies and summarises data using mean				Identifies and summarises data using median				Identifies and summarises data using mode				Best data representation measure			
	P		NP		P		NP		P		NP		P		NP	
	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I
8	X				X				X				X			
9	X				X				X				X			
	2				**2**				**2**				**2**			

Legend: P = Present NP = Not Present A = Appropriate I = Inappropriate

Tina's Percentage of Subject Matter Knowledge of Measures of Central Tendency in Handling Bias

$$= \frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)} \text{ X } 100\% = \frac{\mathbf{8}}{\mathbf{8}} \times \mathbf{100}\% = 100\% \text{ (High)}$$

Level of subject matter knowledge of measures of central tendency in problem solving. In Task 10, Tina carried out the backward mean calculation to obtain the total grade point average for College A based on two semesters as 6.4. She also found the total grade point average for College B based on three semesters as 11.4. Therefore, Tina's knowledge of performing the backward mean calculation to obtain the total from the mean is found to present (P) and correct (A).

Tina explained that the mean which was given for each college was for the number of semesters involved in that particular college. She also added the total grade points of both colleges and the total semesters of both colleges before she proceeded in the weighted mean calculation. Thus, Tina's knowledge related to the representative nature of the mean is found to be present (P) and appropriate (A).

Finally, Tina carried out the forward mean calculation to find the weighted mean. She divided the total grade points of both colleges with the total semesters of both colleges and found the weighted mean as 3.56. Therefore, Tina's knowledge of the forward mean calculation is found to present (P) and appropriate (A) despite the task not explicitly stating to find for the weighted mean.

In Task 11, Tina performed the backward mean calculation and obtained the total hours of TV watched per weekend for both rural and urban groups. She said that she multiplied the number of students in each group with the group average. Tina found the total hours of TV watched for rural group by multiplying 25 with 8. Similarly, for the urban group she multiplied 75 with 4. Thus, Tina's knowledge of performing backward mean calculation was found to be present (P) and appropriate (A).

Tina explained that she multiplied the average of each group with number of students in each group because the given average was for the group. She also mentioned that the totals of both the groups were added and the total number of students of both groups were found before she proceeded with further calculation because she wanted to find the overall average for all the students. Thus, Tina's knowledge related to the representative nature of the mean is found to be present (P) and appropriate (A).

Finally, Tina carried out the forward mean calculation to find the weighted mean. She divided the total number of hours of both groups with the total number of students and found the average as 5. Therefore, Tina's knowledge of the forward mean calculation is found to present (P) and appropriate (A) despite the task not explicitly stating to find for the weighted mean.

Table 4.128

Coding rubrics for determining level of Tina's subject matter knowledge of measures of central tendency in problem solving

Task	Subject Matter Knowledge of Measures of Central Tendency in Problem Solving											
	Backward mean calculation				Representative nature of the mean				Forward mean calculation			
	P		NP		P		NP		P		NP	
	A	I	A	I	A	I	A	I	A	I	A	I
10	X				X				X			
11	X				X				X			
	2				**2**				**2**			

Legend: P = Present NP = Not Present A = Appropriate I = Inappropriate

Tina's Percentage of Subject Matter Knowledge of Measures of Central Tendency in Problem Solving

$$= \frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)} \text{ X } 100\% = \frac{\mathbf{6}}{\mathbf{6}} \times \mathbf{100\%} = 100\% \text{ (High)}$$

Level of subject matter knowledge of measures of central tendency in making inference. In Task 12, Tina calculated the mean scores for both the players by adding all the scores and divided the total with the six set of games that they played. She found that the mean score for Ramu was 19.66 and the mean score for Beng was 21.8. Therefore, Tina's knowledge of summarising equal sized numerical data using the measures of central tendency in this case the mean without the task explicitly stating to do so was found to be present (P) and appropriate (A).

Tina compared the mean scores of both the players and decided on the player that she would recommend to be selected for the team. Since Beng's average was higher than Ramu's, she said that Beng would be recommended. Tina also mentioned that the selection of the player can be done by comparing the lowest and the highest score. However, she mentioned that the average comparison is the best way to compare and decide on the player to be selected because the average considers all the scores for each player. Thus, Tina's knowledge of utilising the appropriate measure of central tendency to compare and draw conclusion on two equal sized numerical data was found to be present (P) and appropriate (A)

In Task 13, Tina summarised the data using the mean or the average. She calculated the average by adding all the scores for each player and then divided the total scores with the number of game sets that they played. She said for Ali she divided his total scores with eight because he played eight game sets. Whereas for Khan, she divided his total scores with six because he played six game sets. Tina found that Ali's average was 22.25 while Khan's average was 23. Therefore, Tina's knowledge of summarising unequal sized numerical data using the measures of central tendency in this case the mean without the task explicitly stating to do so was found to be present (P) and correct (A).

Tina compared the average scores of both the players and decided to recommend Khan because his average score was higher than Ali's. Tina said that the average considered all the game sets that the player played. Tina claimed that the average is the best way to select the player. She added that comparing the lowest or the highest score might not be accurate because it might be just a lucky or an unlucky score for that player. Thus, Tina's knowledge of utilising the appropriate measure of central tendency in this case the mean to compare and draw conclusion on two unequal sized numerical data was found to be present (P) because the mean is the most suitable measure to be used to compare and draw conclusions based on unequal sized data. Tina's knowledge of calculating and drawing conclusions based on the mean was considered to be appropriate too (A).

In Task 14, Tina summarised the data given using the mean or the average. She multiplied the number of students with the number of corrects and added the products. Then she divided the total with the total number of students. Tina calculated the average for 5A as 4.88 and for 5B as 3.8. Therefore, Tina's knowledge of summarising equal sized graphical data using the measures of central tendency in this case the mean without the task explicitly stating to do so was found to be present (P). However, Tina calculated the mean for 5B incorrectly (I).

Tina compared the class averages of both the classes and found that 5A had a higher average than 5B. Therefore, she said that 5A performed better than 5B. Tina mentioned that there are no other ways to determine the class that performed better except the average. She added that the average take count of all the scores for all the students. Thus, Tina's knowledge of utilising the appropriate measure in this case the mean to compare and draw conclusion on two equal sized graphical data was found to be present (P). She explained that the class that had higher average performed better and the average comparison is the best way to determine the class that performed better, so her justification here is considered appropriate (A).

In Task 15, Tina summarised the data given using the mean and the mode. She multiplied the number of students with the number of corrects and added the products. Then she divided the total with the total number of students. Tina calculated the average for 5C as 5.31 and for 5D as 6.19. Tina also identified the mode for 5C as 5 or 6 and the mode for 5D as 7. Therefore, Tina's knowledge of summarising unequal sized graphical data using the measures of central tendency in this case the mean and the mode without the task explicitly stating to do so was found to be present (P). However, Tina calculated the mean for 5C incorrectly which is considered that she might have problem related to the mean calculation when the data was presented in the graphical form (I).

Tina compared the class averages and decided that 5D performed better than 5C because 5D had a higher average compared to 5C. Tina also mentioned that before she calculated the mean, she tried comparing the graphs. However, after realizing that the number of students in both the classes were unequal, she decided to use the mean. She claimed that the mean comparison is the most appropriate way to determine the class that performed better. Thus, Tina's knowledge of utilising the appropriate measure in this case the mean to compare and draw conclusion on two unequal sized graphical data was found to be present (P). She also explained that the class that had higher average performed better, so her justification here is considered appropriate (A).

Table 4.129

Coding rubrics for determining level of Tina's subject matter knowledge of measures of central tendency in making inference

Task	Subject Matter Knowledge of Measures of Central Tendency in Making Inference																															
	Summarises two equal sized numerical data using measures of central tendency				Utilises the appropriate measure to compare and draw conclusion about two equal sized numerical data				Summarises two unequal sized numerical data using measures of central tendency				Utilises the appropriate measure to compare and draw conclusion about two unequal sized numerical data				Summarises two equal sized graphical data using measures of central tendency				Utilises the appropriate measure to compare and draw conclusion about two equal sized graphical data				Summarises two unequal sized graphical data using measures of central tendency				Utilises the appropriate measure to compare and draw conclusion about two unequal sized graphical data			
	P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP	
	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I
Task 12	X				X																											
Task 13									X				X																			
Task 14																		X			X											
Task 15																										X			X			
	1				**1**				**1**				**1**								**1**								**1**			

Legend: P = Present NP = Not Present A = Appropriate I = Inappropriate

Tina's Percentage of Subject Matter Knowledge of Measures of Central Tendency in Making Inference

$$= \frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)} \text{ X } 100\% = \frac{6}{8} \times \mathbf{100}\% = 75\% \text{ (Medium)}$$

Tina's mean percentage of subject matter knowledge for all four constructs $\frac{\mathbf{45.45+100+100+75}}{\mathbf{4}} = \mathbf{80.11}\%$. Therefore, Tina's level of subject matter knowledge of measures of central tendency is high.

Case Study Six: Joan

Joan was 24 years old when she was interviewed. At the point of data collection, she was pursuing a 4-year Bachelor of Science with Education (B.Sc. Ed.) program at a public university. She majored and minored in chemistry and mathematics respectively. Joan obtained A1 in Mathematics and Additional Mathematics in her SPM examination (equivalent to O level examination). She scored a B+ in Mathematics in her Matriculation examination.

At the time of data collection, Joan was in the final semester of fourth year studies. She attained 2.98 in the Cumulative Grade Point Average (CGPA) for her first three years of studies in the public university. She does not have any teaching experience prior to this interview. The following sections discuss the findings of Joan's subject matter knowledge of measures of central tendency in each of the four different constructs, namely with reference to context, in handling bias, in problem solving, and in making inference that were emerged from the clinical interview.

Types of subject matter knowledge.

Subject matter knowledge of measures of central tendency with reference to context.

Task 1. Joan's knowledge of the mean is the average. Joan elaborated on the mean calculation to explain on how the average would have been obtained which clearly indicated her knowledge of the average.

Table 4.130

Excerpt J1

R: How do you think the average was obtained?
S: The average can be obtained by total up all the number of children in Malaysia and divided by the number of families.
R: Can you show and explain using an example?
S: Can I give slightly different example? Let's say family A has 5 children, family B has 6 children and family C has 3 children so to find the average, we total up the number of children and divided by three families, family A, B and C and then we get 14 over 3 and final answer 4.66 but then we consider the average as 4.

R: Why 4?
S: Because 0.66 cannot be counted as children because is incomplete as one.
R: How can the average be 2.3 and not a counting number like 1, 2, 3, or 4?
S: Oh, ok because every family has different number of children and we have to total up the number of children divide by the number of families and this can result in a decimal value.
R: Can you explain further?
S: I would explain that when we total up and divide, we might end up with decimal number. This decimal number might not be logic. Sometimes we can round off this number if to make it more logical to represent children.

Joan mentioned that the average is obtained by dividing the total number of children with the total number of families. She elaborated this in an example where she mentioned that if family A has 5 children, family B has 6 children and family C has 3 children, therefore the average is obtained when the total number of children which is 14 is divided with the total number of families which is 3. She said that the average that she calculated was 4.66 but she considered the average as 4. According to Joan, 0.66 cannot be counted as one complete child therefore she took the average as 4.

Joan explained that since every family has different number of children and the number of children in each of families are totalled up and then divided with the number of families, therefore, she said that we can end up with a decimal value for the average. However, she mentioned that the decimal average might not be logic. Instead she said that this decimal number for the average can be rounded off to make it more logical in representing the number of children. Excerpt J1 describes Joan's points on this.

Task 2. Joan utilised the idea of mode without the task explicitly stating to do so. She mentioned that the mode can represent the shoe sizes because the mode size is favoured by many customers.

Table 4.131

Excerpt J2

R: How would you place the order for the shoe sizes in your shop?
S: I will tabulate all the data which is the purchased shoe size in a table and I look at the highest purchase and then I will order that size more than the others.
R: Can you show me how you did that?
S: I will tabulate my order [explains based on the written answer] I count the numbers for each shoe size and the one with the highest frequency or the highest number of purchase, I will order more.
.
.
.
R: If you want to choose one female shoe size, which shoe size would you choose?
S: I would choose size 5.
R: Why?
S: Because it has the highest demand or the highest purchase.
R: Can the chosen shoe size represent the female shoe size in your shop?
S: Can. Yes.
R: Why?
S: Just like in question (b), it shows that the shoe size is favoured by many customers and has the highest demand.
R: What would you use as the average shoe size?
S: First of all, I have to count all the total shoes purchased and then divide by six shoes sizes and then I get 5.5. So, I will use the average as 5 or 6.
R: Can you explain?
S: I have to find the mean and decide which size will be used as the average. So, I total up everything and divided by six class. So, I got 5.5, so I will choose size 5 or 6.
R: What did you total up?
S: All the shoe sizes times frequencies size purchased. And divide by six class [points at the working in the task sheet].
R: And you said you would choose size 5 or 6 as the average. Why?
S: Because here there is no size 5.5. There is only 5 or 6. So 5.5 is in between 5 or 6.
.
.
.
R: Can we use the mode as the average?
S: No cannot
R: Why cannot?
S: Because we calculate the average and we get 5 point something. And maybe mode depends on highest customers purchased. It not necessarily 5.

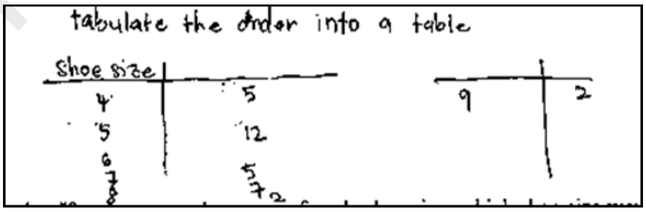


Figure 4.29. Entry J1

Joan mentioned that she would place the order for the shoe sizes in her shop after she tabulated the data. She said that she would place her order more on the size that had the highest purchase. She said that she would count the numbers for each shoe sizes and the one that has the highest frequency or the highest number of purchase, she would order more. This Joan showed clearly in Entry J1.

Joan said that she would choose size 5 if she was given a choice to choose one female shoe size because this particular shoe size has the highest demand or purchase. She also agreed that the chosen shoe size can represent the female shoe size in her shop because of many customers favoured this size which is shown by its highest demand.

However, Joan said that she needed to count the total shoes purchased and then divide with six different shoe sizes when she was asked on what she would use as the average shoe size. She got the average as 5.5 but she said that she would use either 5 or 6. She further elaborated that she would find the mean and then decide on the size that will be used as the average.

Joan mentioned that "total up everything" was referring to the total after all the shoe sizes were multiplied with the respective frequencies. She explained that she chose 5 or 6 as the average because there is no size 5.5 for shoe size and since 5.5 was in between 5 or 6 so it had to be either one of these sizes. Joan said she would not use the mode as the average because the average that was calculated was 5 point something. She said that the mode depended on the highest purchase and is not necessarily 5. The above Excerpt J2 illustrates Joan's points on this.

Task 3. Joan identified the mode based on the highest number of comments in which she believed to indicate the highest frequency. She used the highest number of comments as the value to represent the number of comments made by the students on that day because it was a counting number.

Table 4.132

Excerpt J3

R: What is the mode?
S: Mode is student D
R: How did you obtain this value as the mode?
S: Because student D has the highest frequency or the highest number of comments.
R: If you were to represent one value for the number of comments made by the students on that day, would you use this value?
S: Yes, I would. The number 22 can be used.
R: Why?
S: Because it can be counted.
R: Can it represent?
S: Yes
R: Why?
S: I do not understand the question well. Wait a while. {Looked unsure} It can be any value, right?
R: But the question is asking you whether the value that you got as the mode can be used or not.
S: Yes
R: Why?
S: Because it can be counted.
R: Any other reason behind it?
S: No
.
.
R: Just now the given value is 22 here. Say if this value is now replaced with 6. What is the mode now?
S: Still D
R: How did you obtain the mode?
S: We look at the highest number of comments made.
R: If you were to represent one value to represent the number of comments made on that day, would you use this value?
S: Yes.
R: Why?
S: Ahhh… because the value can be counted just like earlier.
R: Any other explanations?
S: No other explanation.

Joan mentioned that the mode was student D. She said that she based this on the highest frequency or the highest number of comments. Joan said that she would use 22 to represent the number of comments made by students on that day because 22 can be counted. She kept on repeating the same answer when the question was repeated to her several times.

The questions were repeated again after the task was modified where the number of comments for student D was changed from 22 to 6. Joan still mentioned the mode as student D which was based on the highest number of comments. She mentioned that this value can be used to represent the number of comments made by students on that day because it can be counted. The above Excerpt J3 describes this.

Task 4. Joan mentioned that average can mean a "common number". However, Joan's explanation indicated that her knowledge of the average is the mean. Joan also mentioned that the average can be based on the mode but not median. Nevertheless, Joan claimed that the mode cannot be used as the average in this situation because of the information on a group of students watched more than 12 hours of TV per day.

Table 4.133

Excerpt J4

R: What does "average" mean in this sentence?
S: The common number of hours of watching TV per day by the students.
R: How do you think they got this average of 3 hours of TV per day?
S: Average is considered as mean which is we take the data from one population or known as sample. Data on the number of hours TV watched per day from one sample to another sample taken and then total up this. Then divide with the total number of sample which gives the average as 3. So, this one is for one sample [points at the 12 hours of TV per day].
R: Do you think this sample can have an effect on the average?
S: Yes
R: How will it affect the average?
S: Because the number of hours here is far from the mean value given.
R: How do you think that they would have dealt with this information?
S: If it very far, this information should be eliminated.
R: How about here?
S: I think in this situation, they might have added up this information with the group…. the group that watches TV 12 hours per day. They would have added up all together.
R: Why?
S: Because it is considered for Malaysia. It is for one whole Malaysia so they take the information randomly. Because we want to find the average (mean). It is a common number for all. Majority of people watch 3 hours. However, they will also include the 12. They will total up everything and then divide by the number of students.
R: Can you explain using an example?
S: If I do group by group is it ok?
R: It is up to you
S: I use this example, group A, the average is 6 hours per day, group B is 3 hours per day and so on 2 hours per day, 1 hour per day. So, all these hours per day by each group are added and then divide by how many groups there to get the average.

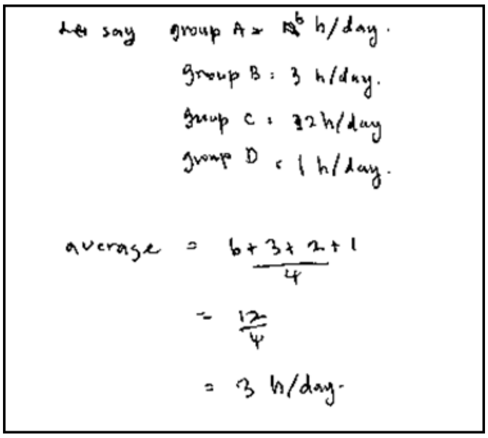


Figure 4.30. Entry J2

According to Joan, average means the common number of hours of TV watched per day. Joan mentioned that the average is considered as mean which is obtained from the data drawn from a population or known as a sample. She elaborated that the data on the number of hours of TV watched per day is collected from one sample to another sample.

After which she said that all the data are totalled up and then divided with the total number of samples. Joan points out to the information that a group of students watched more than 12 hours of TV per day as "this one is for one sample". She mentioned that the information can have an effect on the average because the number of hours in the information is far from the given mean value. She also said that this information should be eliminated because it is far from the mean.

However, Joan later mentioned that in this situation, the information might have been added up along with the rest of the information on the number of hours TV watched. She said that this is because the average is considered for the whole Malaysia. She also added that the information was taken randomly and because the average is considered which means that the average is a common number for all. Joan said that the majority would have watched 3 hours. However, she also said that 12 hours would have been included in the average calculation. She mentioned that all the information regardless how many hours watched would have been totalled up and then divided with the number of students.

Joan elaborated this further using an example dealing with groups that had different group averages for the number of hours of TV watched per day. She said if group A has an average of 6 hours per day, group B has 3 hours per day, and so on 2 hours per day and 1 hour per day. Thus, she said that the average is calculated by all the group averages are added up and then the total is divided with the number of groups as shown in Entry J2. The above Excerpt J4 illustrates Joan's points on this.

Table 4.134

Excerpt J5

R: Besides the way that you have showed and explained, do you think is there any other ways to find the average?
S: No.
R: Can median or mode be used as average?
S: Cannot
R: Why?
S: Can. Sorry
R: Why?
S: A while please. I think can. I draw a graph and explain. Let's say we have a graph. This is the distribution. The one as the highest is the mode. If look at the normal distribution, this is the mean. So, we can still find the average using mode.
R: What about the median?
S: Median is the middle class. Not median. Median cannot I think. {Looked unsure}.
R: How about in this question, do you think the mode is used as the average?
S: No
R: Why?
S: Hey can I think because this is for one group only, there might be other groups that watch 3 hours of TV per day so maybe that is the mode class {looked unsure}.
R: How about here is the mode appropriate to be used as the average?
S: No.
R: Why?
S: Because the distance between the mode and 12 is far.
R: So how?
S: We find the mean. {Looked very unsure of her answer}
R: What about median?
S: Median for average. No. Still no.
R: So how do you think the average was obtained here?
S: Like I said just now by finding the mean. So, I can {} the number of group and then I sum up everything and divide.

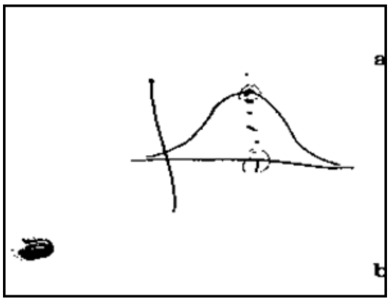


Figure 4.31. Entry J3

Joan mentioned that she was not aware of any other ways to find the average beside the one that she elaborated. When she was asked if the median or the mode can be used as the average, she said no. However, Joan changed her mind and said that maybe can be used after she was probed. She requested for some time to think. She sketched a graph on the task sheet as shown in Entry J3. Joan explained based on the normal distribution graph that she sketched. She said for the normal distribution, this is the mean pointing at the mid of the graph and the highest peak is the mode. Therefore, she said that the average can be based on the mode. However, Joan mentioned that the median which is the middle class cannot be used. Nevertheless, her response indicated some uncertainty.

Although Joan earlier mentioned that the average can be based on the mode but her subsequent responses showed that she was not very certain about this. First, she said no for the use of the mode as the average in this situation but later she said can because she said the information on a group of students watched more than 12 hours per day is just one group. She elaborated that there might be other groups that watched 3 hours so this can be the mode class.

However, when Joan was probed further, she claimed that the mode is not appropriate to be used as the average in the given situation. She explained that this is because the distance between the mode and 12 is far. Instead she said she would find the mean. However, Joan sounded a bit unsure of her answer. She also held on to her answer earlier that the median cannot be used as the average. Joan said that the average can be found by calculating the mean in the similar way that she described earlier using the group averages. The above Excerpt J5 describes Joan's points on this.

Task 5. Joan identified the median by arranging the data in ascending order first. For odd numbered data, the median is the middle number in the arrangement. On the other hand, for even numbered data, the median is the mean of the two middle numbers in the arrangement. Joan mentioned that the mean is not adequate in representing the weight of all ten children whereas she mentioned that the median is more appropriate.

Table 4.135

Excerpt J6

R: Which is the weight for the median child?
S: 19
R: How did you find that?
S: I arrange everything in ascending order, there are nine and I take the middle value
R: Which is the median if we include another child that weighs 43 kg?
S: Still the same. I just add 43 kg at the back and then there are 10. To find the median I have to consider two numbers in between. I plus these two numbers and then divide by two.
R: Is there any other way to find the median?
S: No. This is the only way that I know.
R: Is it adequate to use the mean to represent the weight of the 10 children?
S: No.
R: Why?
S: Because from the calculation, I obtain the mean as 22.2 which is quite different from the weight that are given here. So, does not look very accurate.
R: So, what do you think will be more appropriate to use to represent the weight of the 10 children?
S: Median
R: Why?
S: Because it lies in between the lightest and the heaviest.
.
.
.
R: How about if mean is used in this situation?
S: Here in this situation, the 10th weight will not be used because it differs a lot. I would eliminate the value. I would find the mean based on 9 children.
R: Can this mean be used to represent the weight of the children?
S: Can.
R: Why?
S: Because the value of this mean which I calculated as 19.8 is not that far away from each of the other values.

Joan identified the median as 19. She explained that she would arrange the data in ascending order first. She mentioned that since there are nine data, she would take the middle data in the arrangement as the median. Joan also identified the median as 19 after 43 kg was included. She said that she added 43 kg at the back of the arrangement which made the total data in the arrangement as 10. She mentioned that to find the median, she considered the two middle numbers. She said that she added these two numbers and then divided by two to get the median as 19. Joan also mentioned that there are no other ways to obtain the median.

Joan said that the mean is not adequate to represent the weight of all ten children. According to her mean calculation, she got the mean as 22.2 which she said to be quite different from the other weights given in the data. She also claimed that this mean did not look very accurate. Instead, Joan claimed that the median is more appropriate to represent the weight of all ten children because the median weight fell in between the lightest and the heaviest weights in the data.

However, Joan explained if the mean is still considered in this situation, then the weight of the tenth child 43 kg will not be used. She said that she would eliminate this weight and find the mean based on the remaining nine weights. Joan explained that the new mean that she calculated was 19.8 which is not far away from the other weights, therefore it can represent the weight of the children. The above Excerpt J6 illustrates Joan's point on this.

Task 6 and task 7. Joan placed the prices for all seven bags of crisps randomly. She utilised the total price that she had calculated from the given average as a guide. Joan had imagined that her choice of prices should follow the normal distribution bell shape in which she claimed is commonly used in statistics.

Table 4.136

Excerpt J7

R: How did you place the price?
S: I put all 27.
R: Why?

S: Because to find average, I have to sum up everything and divided by how many crisps we have. So, 27 divide by … ahhh sum up everything and divide by 7, I get back 27. R: Can explain by showing me how you did that? S: Ok. So, 27 because there are 7 bags so times 7 divide by 7 so the average is 27. R: Any other reason? S: No. Just that. It is easy to put 27 for all

Joan placed the price for all seven bags of crisps as 27. She explained that in order to find the average, all the prices have to be summed up and then divided with 7 in order to get back the average as 27. She elaborated that 27 times with 7 and then divided again with 7 would give the average as 27. Joan said she had no other reason for placing 27 for all the bags of crisps. She just felt it was easy to do so. The above Excerpt J7 illustrates Joan's points on this.

Table 4.137

Excerpt J8

R: How did you base these values? S: I just put any values here and I add up everything. Same technique divides by 7 to obtain 27. R: So how did you base each value here? S: Since the average is 27, I multiply with 7. Total that I get 189. R: How did you base these values? S: Just randomly put actually. Randomly take any value. R: When I look at the values that you have taken, I can see that you have taken big numbers and also small numbers. Why? S: Yes, I just imagine a shape that starts with a small number followed by big ones. That is how I chose the values too. R: Why the shape? S: I cannot explain this. I just image a normal distribution bell shape R: Can you explain again, how you got these values here? S: Ok. There are 7 crisps. In statistics, the bell shape is used and the data will be starting from small to big values. Then I just put some values when I total up everything must be equal to 189. Just like that only.

Joan mentioned that she placed any values for the prices of the bags of crisps in Task 7. She said that similar to what she did in Task 6, she added up all the prices and then she divided with 7 to get 27. She explained that since the average was 27, she multiplied the average with 7 and got the total as 189. She based on the total and randomly placed any value.

Joan explained that she imagined the data followed a particular shape that started with small numbers to big numbers. Therefore, she chose the prices based on this idea too. However, when she was probed to explain why she imagined such a shape, she first mentioned that she could not explain but later she said that she imaged a normal distribution bell shape.

She further elaborated that the bell shape is commonly used in statistics and the data starts from small to big numbers. Therefore, Joan said she followed the similar idea to base the values for the prices of the bags of crisps. She said she made sure that when she totalled up everything (referred to the prices) she got the total as 189. However, she mentioned that she placed the prices randomly. She just made sure that when she totalled up all the prices that she gets 189. Excerpt J8 describes Joan's points on this.

Summary.

Table 4.138

Subject matter knowledge with reference to context

Task	Statistical element	Conclusion
Task 1	Mean as average	Joan's knowledge of the mean is the average. She elaborated that the average is obtained when the total number of children is divided with the total number of families. Joan also explained on the mean calculation in her example. Joan said that a decimal average for children can be rounded off to a whole number because the decimal is incomplete as one. Joan said that in the mean calculation involving the division of the totals, a decimal value is possible. However, she mentioned that this decimal value might not be logic given to the context of the data. She explained that to make it more logical, the decimal average can be rounded off so that it represents the context of the data.
Task 2	Idea of mode	Joan utilised the idea of mode without the task explicitly stating to do so. She used the idea of frequency tally and determined the mode based on the highest purchase or the highest number of frequency.
	Mode as data representation	Joan said that the mode can be used to represent the shoe sizes in her shop. She said that this is because the mode size is favoured by her customers and has the highest demand.
	Mode as average	Joan explained on the calculation of the mean when she was asked on what she would use as the average shoe size. She said that she totalled up all the shoe sizes and then divided the total with six different shoe sizes. Joan got the shoe size as a decimal number in which she said

		that she would round off to a whole number. Joan said that the mode size cannot be used as the average size. She said that this is because the mode is based on the highest purchase whereas the average is calculated. It happened to be that in this situation, the average calculated can be rounded off to the mode size. However, it does not always happen in such way that the values coincide all the time.
Task 3	Idea of mode	Joan obtained the mode based on the highest number of comments shown in table in which she referred as the highest frequency. Her knowledge of the mode is categorical.
	Mode as data representation	Joan said that the mode can be used to represent the data because the mode can be counted.
Task 4	Median as average	Joan said that the average referred to the "common number of hours" students watched TV per day. However, her explanation on how the average was obtained showed that her knowledge of the average is the mean. She elaborated on the mean calculation in her example too. Joan said that the information on the extreme data does have an effect on the average. Although Joan knew that any extreme data can be excluded from the mean calculation but for this situation she said that this information was added into the mean calculation. According to her, the average is for one whole Malaysia so it must consider all the information including the extreme. On the other hand, Joan said that the average can be based on the mode but she claimed that the mode is not suitable for this situation because of the gap between the mode and the extreme data is quite big. She also said that the median is not suitable to be used as the average. Therefore, Joan mentioned that the mean is still the best to be used as the average because it considers all the data including the extreme.
Task 5	Idea of median	Joan carried out the median procedure by arranging the data in ascending order first. She took the middle number as the median for the case where the data is odd numbered. For the case where the data is even numbered, she said that the median is obtained after the two middle numbers are added and then divided by two.
	Robustness of measures	Joan said that the mean is not adequate to represent the weight of all ten children. She said that from her calculation of the mean, she obtained the mean as 22.2 which she said is quite different than the rest of the data. Instead, she said that the median is more appropriate to be used because the median is in between the lightest and the heaviest weight. However, Joan said that the mean calculated after excluding the extreme weight can be used to represent the weight of the children because this mean value is around the rest of the data.

Task 6	Idea of average	Joan said that she placed the prices for all the bags of crisps as 27. She explained that when she summed up this and divided with 7, she got the average as 27. Joan elaborated on the backward mean calculation when she said that 27 times with 7 and then divided with 7 again would give her 27. Therefore, she said that it was easy for her to put all 27.
Task 7	Idea of average	Joan said that she placed randomly the values for the bags of crisps as long as all the seven values once added up and then divided with 7, gives her 27. Joan also utilised the backward mean calculation to get the total price. She said that she imagined the values that she placed randomly as the numbers that fitted a normal distribution bell shape which starts with small numbers to big numbers. She mentioned that although these numbers are in random but she just made sure when totalled up it equalled to the total price.

Subject matter knowledge of measures of central tendency in handling bias.

Task 8 and task 9. Joan identified the measures of central tendency and elaborated on the procedures to obtain these measures. Joan claimed that the mean is the best measure to represent a set of data in the presences of an error. According to her, only in the mean, the error can be excluded which is not possible in the median or the mode.

Table 4.139

Excerpt J9

R: What are the measures of central tendency that you can obtain based on the above data?
S: All three can.
R: Can you mention to me, what are the three?
S: Mean, median and mode
R: How did you obtain these measures?
S: To find the mean as usual. Mean is adding up all the values and then divide by the number of data. If I want accurate I will eliminate 15.3. I sum up everything excluding 15.3 and divide by eight.
R: Why did you exclude 15.3?
S: Because it is too far from the other data. So, I eliminate it and divide by eight. So, I get 6.17.
R: How did you get the median?
S: Median, I arrange it according to the ascending order and then I pick the middle value which is 6.2.
R: Any other ways of finding the median?
S: No only this way.
R: How did you get the mode?
S: I would find the frequency of occurrence of each mass. 6.3 has the highest frequency which is 3. So, 6.3 is the mode
R: Which is the best measure of central tendency to represent the actual weight of this object?
S: Mean
R: Why?

S: Because in mean, we can eliminate any unnecessary data that we do not want but, in the mode and the median, this value still exists.
R: But do think this value affects the median or the mode?
S: Yes, it does affect the median or the mode.
.
.
.
R: Why not mode?
S: Mode only represent the most frequent and not the overall.
R: Why not median?
S: Median considers the middle number only and not the overall. So, the best is the mean because it considers all data.
R: But in mean you did not consider 15.3?
S: Yes, because it is far, we can eliminate this value for better accuracy. This value can be considered error.

Joan mentioned that all three measures of central tendency, the mean, the mode and the median can be obtained from the data given in Task 8. She explained that the mean is calculated by adding up all the values given in the data and then divide by the total number of data. However, Joan mentioned that if she wanted the mean to be accurate than she would eliminate 15.3 and base the calculation on the remaining eight data. She said that 15.3 is far from the other data, therefore she eliminated it. She got the average as 6.17.

Joan explained that to find the median, she would arrange the data in ascending order first and then she would pick the middle number in the arrangement which is 6.2 as the median. However, she mentioned that she is not aware of any other ways to find the median. Joan added that in order to find the mode, she would find the frequency of occurrence for each mass. She found that 6.3 had the highest frequency which is 3, thus she mentioned that 6.3 is the mode.

According to Joan, the best measure of central tendency to represent the actual weight of this object was the mean. She explained that in the mean, unnecessary data that we do not want can be eliminated but this cannot be done in the median or the mode. She also mentioned that the existence of this unnecessary data would affect the median or the mode.

For Joan, the mode represented the most frequent whereas the median represented the middle number. She said that these two measures do not represent the overall, unlike the mean that considers all the data. However, Joan mentioned that in this situation 15.3 was considered as an error because it is far from the other data in which can be eliminated for better accuracy. The above Excerpt J9 illustrates Joan's points on this.

Table 4.140

Excerpt J10

R: What are the measures of central tendency that you can obtain based on the above data?
S: Only two which is the mean and median.
R: How did you find the mean?
S: Mean which I sum up everything without considering 180 000 and then divide by 4.
R: Why?
S: Just like the earlier question, this value is very far from the rest and I consider as an error. I only take four data and divided by four, I get 51 000.
R: How about the median?
S: Median, I just arrange it according to ascending order and I pick the middle one, so I get 54 000. And there is no mode because each salary only has one frequency.
R: Which is the best measure of central tendency to represent the actual salary of the company?
S: Mean
R: Why mean?
S: Because if median, it still includes the error. But mean can exclude the error to get better accuracy.

Joan obtained only two measures of central tendency which are the mean and the median based on the data given in Task 9. She calculated the mean by adding all the data except 180 000 and then she divided the total with 4. She explained that just like in the earlier task, 180 000 was far from the rest of the data in which she considered as an error. Therefore, she added the remaining four data and divided the total with four to get the mean as 51 000.

Joan got the median as 54 000. She explained that she first arranged the data in ascending order and then she picked the middle value as the median. According to Joan for this situation there is no mode. She said this was because each salary had only a frequency of one.

For Joan, the mean is the best measure of central tendency to represent the actual salary of the company. She said that in the mean, the error can be excluded for better accuracy which she claimed cannot be done in the median. The above Excerpt J10 describes Joan's points on this.

Summary.

Table 4.141

Subject matter knowledge in handling bias

Task	Statistical element	Conclusion
Task 8	Identifies and summarises data using mean	Joan found the mean by adding eight data after excluding the extreme data and then divided the total with eight. She said that this mean is a more accurate mean that the mean found by adding all the nine data and then divide the total with nine.
	Identifies and summarises data using median	Joan identified the median by first arranging all the data in ascending order and then she located the middle value as the median.
	Identifies and summarises data using mode	Joan identified the mode based on the highest frequency. The data point that occurred the most is considered as the one with the highest frequency.
	Best data representation measure	Joan mentioned that the mean is the best representation measure. She said that the mean that she obtained had excluded the unnecessary data but this is not the case for the median and the mode. According to her the median and the mode is still affected by this extreme data. Joan also mentioned that the mean considers all the data which is not the case for the mode that considers only the most frequent and the median which considers only the middle value. She also said that the mode and the median do not consider the overall like the mean.
Task 9	Identifies and summarises data using mean	Joan calculated the mean by adding four data values after excluding the extreme data and then the total was divided with four.
	Identifies and summarises data using median	Joan identified the median by first arranging all the data in ascending order and then she located the middle value as the median.
	Identifies and summarises data using mode	Joan said there is no mode for this situation because all the data occurs only once which means that each data had only one frequency and no highest frequency.
	Best data representation measure	Joan mentioned that the mean is the best measure to represent the data because it was calculated after excluding the extreme data. According to her, the median still included the extreme data.

Subject matter knowledge of measures of central tendency in problem solving.

Task 10. Joan mentioned that the given GPA for each college is the average of the grade points for all the semesters involved for that college. Based on this Joan found the overall weighted mean.

Table 4.142

Excerpt J11

R: What was the student's overall GPA?
S: [Shows the answer] I get 3.56
R: How did you get this?
S: Let's say, college A two semesters and earned 3.2 GPA. This one is the mean/ average for college A. So, there are two semesters, I consider as 3.2 plus 3.2 divide by two still get 3.2. So, for this one is the same also, for three semesters, 3.8 plus 3.8 plus 3.8 divide by three semesters you get 3.8. So, to obtain the overall mean, I have to consider those. 3.2 times 2 plus 3.8 times 3 and divide by total semesters and I get 3.56.
R: Why times two here?
S: Because there are two semesters
R: What about times three there?
S: Because there are three semesters for college B
R: Now why are you dividing by five?
S: Because I want to take all the semesters that they have studied. I want to find the overall GPA for both colleges and all semesters involved.

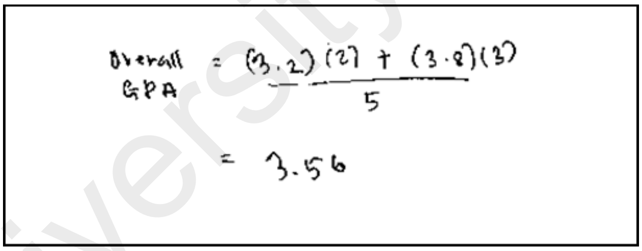


Figure 4.32. Entry J4

Joan mentioned that she got the overall GPA as 3.56 in which she wrote clearly her overall GPA calculation in the task sheet as shown in Entry J4. She mentioned that the GPA for college A was 3.2 which was for two semesters. She explained that when she added 3.2 with 3.2 and then divided with two, she got 3.2. She mentioned that the same applied to college B. She said that since there are three semesters involved, so when she added 3.8 with 3.8 and with another 3.8 then divided the total with three, she got 3.8. Joan said that to find the overall GPA, she considered all this. She said that she multiplied 3.2 with two and added this with 3.8 multiplied with three. Then she said that she divided the total with the total semesters involved and she got the overall GPA as 3.56.

Joan elaborated clearly that the reason she multiplied the GPA for college A with two was because of a total of two semesters were involved in college A. Similarly, she said that the GPA for college B was multiplied with three because a total of three semesters were involved in college B. Finally, she also mentioned that the reason she divided with five was because she wanted to consider for all the semesters the student studied. She said that she wanted to find the overall GPA based on both colleges and all semesters involved. The above Excerpt J11 describes Joan's points on this.

Task 11. Joan explained that each group had its own average for the group. Utilising this information, Joan found the total for each group. She later combined the totals of each group and utilised the overall totals in the weighted mean calculation.

Table 4.143

Excerpt J12

R: Show clearly the average TV viewing time per weekend for all 100 students.
S: So, to find the average for all the100 students, there are 25 students who watched 8 hours per weekend. Since it is mentioned here as the average, so 25 times 8 divide by 8, I still get 25. Here, I divide by 25, I still get 8 hours. Same goes with this, 75 watch 4 hours. So, to find the average for 100 students. 25 times 8 plus 75 times 4 and then divide 100.
R: Why 25 times 8?
S: Because the average 8 is for 25 students. I want to find the total for this group.
R: Why 75 times 4?
S: Because the average 4 is for 75 students. I want to find the total for this group.
R: Why 100?
S: For all students, the total students.
R: Why are you adding this two?
S: Because I find the total hours of TV watched by each group. I total up the total for both groups. Which will give me the overall total hours of TV watched by all students.
R: What are you finding here?
S: Mean actually.

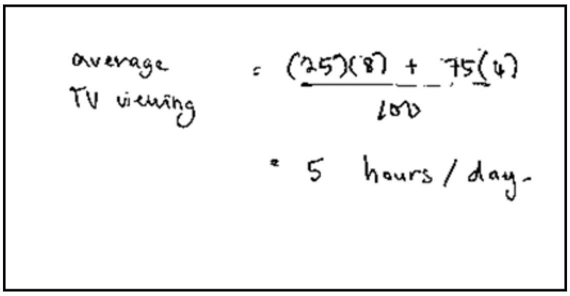


Figure 4.33. Entry J5

Joan showed clearly the calculation that she used to find the average TV viewing time per weekend for all 100 students. She elaborated that 25 rural students watched 8 hours of TV per weekend and 75 urban students watched 4 hours of TV. She said that she multiplied 25 with 8 because the given average 8 was for 25 students and that she wanted to find the total TV viewing time for this rural group. Similarly, she multiplied 75 with 4 because the given average was for 75 students and that she wanted to find the total TV viewing time for this urban group. This is shown in Entry J5.

She said that she added both the totals of the groups that gave her the overall total TV viewing time for all 100 students. After which she divided with 100 to give her the average TV viewing time for all 100 students. Joan mentioned that she was actually finding the mean here. The above Excerpt J12 illustrates Joan's points on this.

Summary.

Table 4.144

Subject matter knowledge in problem solving

Task	Statistical element	Conclusion
Task 10	Backward mean calculation	Joan solved for the weighted mean for Task 10 without the task explicitly stating to do so. She had carried out the backward mean calculation, where she obtained the total for each group of data set from the given group means.
	Representative nature of the mean	Joan used the knowledge related to the representative nature of the mean in making sense of why data sets need to be combined in the weighted mean.

	Forward mean calculation	Joan also carried out the "forward" mean calculation to calculate the weighted mean.
Task 11	Backward mean calculation	Joan solved for the weighted mean for Task 11 without the task explicitly stating to do so. She had carried out the "backward" mean calculation, where she obtained the total for each group of data set from the given group means.
	Representative nature of the mean	Joan used the knowledge related to the representative nature of the mean in making sense of why data sets need to be combined in the weighted mean.
	Forward mean calculation	Joan also carried out the "forward" mean calculation to calculate the weighted mean.

Subject matter knowledge of measures of central tendency in making inference.

Task 12. Joan mentioned that she would select the player based on the average score and the median. However, after she got the median for both the players, she changed her mind and said that the median is not suitable to be used. She said that this was because the median for both the players are almost the same. Therefore, Joan stated that she would select the player based on the mean because the mean considered the overall data.

Table 4.145

Excerpt J13

R: Who would you recommend him select?
S: Beng.
R: Why?
S: Because the average value score is higher than Ramu.
R: How?
S: Based on the mean calculation
R: Is there any other way that you can use to choose the player?
S: If I look at the median [thinks for a while] {looked unsure}.
R: How?
S: [Finds the median for Beng] The median for Beng I get 21.5. Median for Ramu is 20.5.
R: Who would you choose?
S: I would choose Beng
R: Why?
S: No. I cannot choose based on the median because the median is almost the same. No. I do not prefer median.
R: Why?

S: I prefer mean
R: Is there any other ways that you can select the player?
S: Consistency
R: How?
S: That one I am not so sure, maybe look at the range between the scores. This is difficult to be used. So, I prefer the mean.
R: Why?
S: Because it sees the average score.
.
.
.
R: Why you think average is more suitable to use to choose the player?
S: Because it takes count the overall data. We know that the average score can predict the next score.

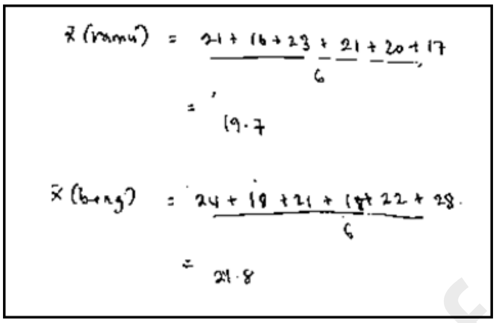


Figure 4.34. Entry J6

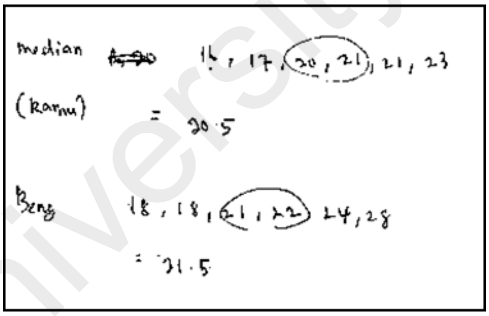


Figure 4.35. Entry J7

Joan mentioned that she would recommend to select Beng. She based her decision on the average score in which she mentioned that Beng's average score is higher than Ramu's. According to Joan, the average score was obtained from the mean calculation which is shown in Entry J6.

Joan said that the median can also be used to choose the player but she looked rather unsure about this. She carried out the procedure to find the median for both Beng and Ramu which is shown in Entry J7. She found that the median for Beng was 21.5 higher than the median for Ramu which was 20.5. Therefore, she mentioned that she would choose Beng. However, later she changed her mind by saying that the median cannot be used to select the player because both the players had almost the same median. She said that she would preferred the mean.

Joan said that another way to determine the player to be selected is by looking at the consistency of the scores. However, she was not so sure on how to look at it. She said maybe by looking at the range but she claimed that it is difficult to use consistency of the scores to select the player. Therefore, she said that she preferred the mean because it is the average score. For Joan, the mean or the average score takes into account the overall data and she said that it can be used to predict the next score. Excerpt J13 describes Joan's points on this.

Task 13. Similar to Task 12, Joan initially mentioned that she would select the player based on the mean and the median because the values for the measures that she obtained for both the players were almost the same. However, Joan changed her mind after she was probed on the different number of scores involved for both the players. After when she decided that she would utilise the mean to select the player because she considered the mean as fair.

Table 4.146

Excerpt J14

R: Who would he select from this group below and why?
S: I will use the mean and the median. Khan will be selected since his mean is higher than Ali's. Same goes to the median.
R: Are you using median to select the player here?
S: Yes
R: Why?
S: Because the value of median and mean obtained is almost the same and only slightly different.
R: If look at the data, Ali got 8 scores and Khan has 6 scores. How about this?
S: Yes. The number of trials is different.

R: How would you determine the player to be selected?
S: I would choose based on the mean.
.
.
.
R: You mentioned mean and median. So how would you select the player?
S: I would prefer using the mean.
R: Why?
S: Because the mean takes the average.
R: Can you explain further?
S: Because the mean is considered fair.
R: Why?
S: Like if we see maybe 1 game he scored the next he doesn't and so. So we take the overall by considering all the games that he won and lose by finding the mean.

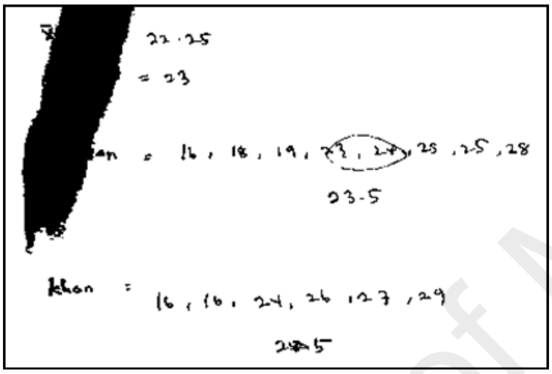


Figure 4.36. Entry J9

For Task 13, Joan mentioned that she would use the mean and the median to select the player. She said that Khan would be selected because his mean and his median are higher than Ali's. She explained that she utilised both the mean and the median because both of the values were almost the same. However, after she was probed on the different number of scores for both the players where Ali had eight scores while Khan had only six scores, Joan decided that she would choose the player based on the mean. She said that she preferred the mean because the mean takes the average. She also claimed that the mean was considered fair. She added that the mean takes the overall by considering all the games the players played regardless whether they won or lost. The above Excerpt J14 and Entry J9 describes Joan's points on this.

Task 14. Joan had a few methods that she used to determine the class that performed better. She said that the class that performed better can be based on the shape of the graphs, the consistency of the scores or even the mode. However, Joan did not utilise the mean because the mean that she calculated for both classes were apparently the same.

Table 4.147

Excerpt J15

R: Based on the graphs below, which class did better?
S: 5B
R: Why?
S: Because of the distribution.
R: How?
S: [Paused a while] {looked unsure} I be frank here. 5B has almost a bell shape which is like a normal distribution. This is preferred as compared to the above graph. But this graph {referring to 5A} is not consistent.
R: But the question asks you which class did better, referring to the performance of the class?
S: But if we look at consistency than it is 5A.
R: Why?
S: Because it increases and consistent. While this {referring to 5B} the graph is going up and down.
R: Besides this, is there any other way to determine which class did better?
S: We can look at the mode, 5B did better because the mode is 4 but 5A got no mode. So based on the mode 5B performed better.
R: Anything else?
S: I have calculated the mean but it gives same value so I cannot determine which class did better.
R: How did you calculate the mean?
S: I just count the number of students and divided by 4, 4 questions.
R: This is for 5A. How did you find for 5B?
S: The same.
R: Can you write down and show me how you found the mean?
S: (She writes down the calculation for mean on the task sheet)
R: So, which class did better?
S: Since both mean values are the same, none of them did better. They are equal. So, we cannot utilise the mean to decide.
R: So how?
S: I would use the mode and the shape of the graph. 5B did better.

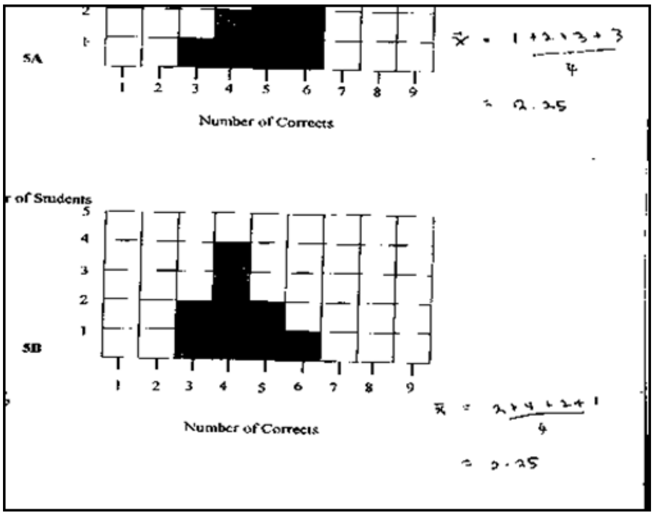


Figure 4.37. Entry J8

Joan mentioned that 5B performed better than 5A. She explained that she based her decision on the distribution. However, she looked a bit unsure of her answer. Joan explained that 5B had an almost bell shape, the shape of a normal distribution which she claimed was preferred as compared to the graph of 5A which she said was not consistent.

Nevertheless, Joan also compared the "consistency" of the classes. Based on this, she claimed now that 5A performed better. She explained that the graph for 5A is seen as "increases" and "consistent" as compared to 5B which showed an up-down trend.

Joan also mentioned that another way to determine the class that performed better would be based on the mode. Here, she claimed that 5B did better because 5B's mode was 4 whereas 5A had no mode. She added that although she had calculated the mean but she cannot utilise the mean to determine the class that performed better because both the classes had the same mean value.

Joan explained that she calculated the mean by dividing the number of students with the number of questions which is shown in Entry J8. Finally, Joan said that she would decide based on the mode and the shape of the graph and she claimed that 5B performed better. The above Excerpt J15 illustrates Joan's points on this.

Task 15. Joan claimed that the performance of the classes cannot be determined if the number of students for both the classes are not the same. However, when she was told that she still has to choose one class that performed better, she made her decision based on the average.

Table 4.148

Excerpt J16

R: Based on the graphs below, which class did better?
S: None
R: Why?
S: Because the number of students is not the same. So, we cannot determine which class did better and which not.
R: But if you still have to choose one class that did better?
S: I will choose 5C.
R: Why?
S: Because the average number of corrects is higher than 5D.
R: How?
S: I just count everything here. For 5C, 1 plus 4 plus 6 plus 7 plus 7 plus 6 plus 4 plus 1. I got 36 divide by 8, I got 4.5.
R: What about 5D?
S: 1 plus 1plus 2 plus 2 plus 4 plus 6 plus 4 plus 1. I got 21 divide by 8, I got 2.6 something.
R: So, which class did better?
S: 5C since the mean value is 4.5.
R: Besides using average, is there any other way to find which class did better?
S: No. Only this. Here I prefer to use the average. Nothing else. The number of students is different for both classes.

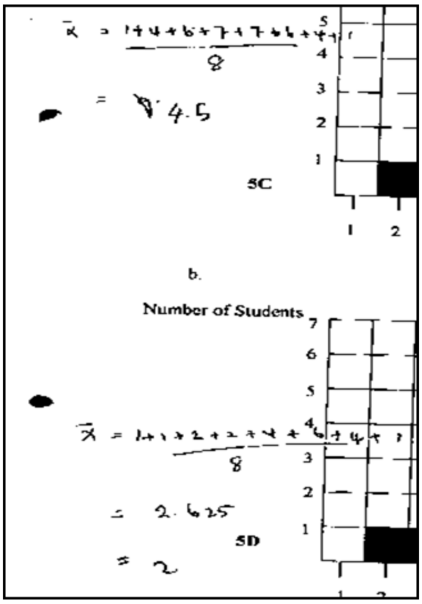


Figure 4.38. Entry J10

Joan first mentioned that none of the class did better. She explained that this was because the number of students were not the same for both classes. Therefore, she said that she cannot determine which class did better and which not. However, when Joan was asked to choose one class that performed better, she chose 5C. She based her decision on the average number of corrects and that 5C had higher average than 5D.

Joan explained that she found the average by adding all the number of students and then divided with eight for both classes. She got the average for 5C as 4.5 from the division of 36 with 8 and the average for 5D as 2.6 something from the division of 21 with 8 which can be seen in Entry J10.

She said that since 5C had a mean value of 4.5 which is higher than 5D, so this class performed better. Joan said that there are no other ways to determine the class that performed better and that she preferred to use the average. She also said that this was because the number of students involved for both the classes were different. The above Excerpt J16 describes Joan's points on this.

Summary.

Table 4.149

Subject matter knowledge in making inference

Task	Statistical element	Conclusion
Task 12	Summarises equal sized numerical data using measures of central tendency	Joan summarised equal sized numerical data using the mean and the median.
	Utilises the appropriate measure to compare and draw conclusion about two equal sized numerical data	Joan said that she would utilise the average or the mean to compare and decide on the player. Joan mentioned that the median can also be used to select the player but she realized that both the players had almost the same median therefore she said she would not prefer to use the median here. Joan explained that she preferred the mean because it takes into account the overall data.
Task 13	Summarises unequal sized numerical data using measures of central tendency	Joan summarised two unequal sized numerical data using the mean and the median.
	Utilises the appropriate measure to compare and draw conclusion about two unequal sized numerical data	Joan initially said that she would select the player based on the mean and the median. She said that the values of the mean and the median for both players were almost the same. However, after she realized that the data are of unequal size, Joan decided that she would base her decision on the mean. She considered the mean to be fair because it considers all the games the player had played regardless whether he won or lost.
Task 14	Summarises equal sized graphical data using measures of central tendency	Joan summarised equal sized graphical data using the mode.
	Utilises the appropriate measure to compare and draw conclusion about equal sized graphical data	Joan had a few ways to compare and decide on the class that performed better. She said that she based on the shape of the graphs, on the consistency of the scores and the mode. According to her, if the graph had a normal distribution bell shape than it is considered that the class had better performance. She also compared the consistency of the scores. However, Joan mentioned that she would decide using the mode and the shape of the graphs. She said that she does not prefer to use the mean because the mean she calculated showed both classes had the same mean. Joan calculated the mean by dividing the total number of students with the total number of questions which she said was equal to four.

Task 15	Summarises unequal sized graphical data using measures of central tendency	Joan summarised two unequal sized graphical data using the mean.
	Utilises the appropriate measure to compare and draw conclusion about unequal sized graphical data	Joan initially said that she cannot determine the class that performed better because both classes had unequal sized data. However, when she was insisted to decide, she utilised the average or the mean. Joan calculated the mean by dividing the total number of students with the total number of questions which she said to be eight in this situation.

Level of subject matter knowledge of measures of central tendency.

Level of subject matter knowledge of measures of central tendency with reference to context. In Task 1, Joan mentioned that the average would have been obtained from the division of the total number of children in Malaysia with the total number of families. She also showed the mean calculation in her example. In her example, Joan rounded off the decimal average that she got to a whole number because she said that the decimal part is incomplete as one. She said that the decimal average is not logic given to the context of the data but it can happen as a result of any division. She added that to make the average more logical to represent the context of the data, it can be rounded off. Therefore, Joan had presented the knowledge of the mean as the average (P). Her justification about the mean as the average is considered appropriate (A).

In Task 2, Joan utilised the idea of mode without the task explicitly stating so. She carried out the frequency tally and tabulated the data. She identified the mode as 5 based on the highest frequency or the highest number of purchase. Therefore, Joan's idea of the mode was found present (P) and correct (A).

Joan said that the mode can be used as a form of data representation. She explained that this was because the mode size is favoured by many customers and had the highest demand. Therefore, Joan's knowledge of the mode as a form of data representation was found to be present (P) and appropriate (A).

However, when Joan was asked on what she would use as the average shoe size, she said that she had to calculate. Joan calculated the shoe size as the total shoe sizes purchased divided by six shoe sizes. She got the average as 5.5 but she said that she would use either 5 or 6 because only size 5 or 6 is available. Joan said that she found the mean here. She also said that the mode cannot be used as the average because it is based on the highest customer purchase. Therefore, Joan's knowledge of the mode as the average was not only found to be not present (NP) but also problematic because she lacked in the knowledge that the mode can be used as the average (I).

In Task 3, Joan identified the mode as student D. She based her decision that student D had the highest frequency or the highest number of comments. Joan identified the mode incorrectly based on the highest number of comments which she thought is indicating the highest frequency. Therefore, Joan's idea of the mode was found to be not present (NP) and incorrect (I).

Joan said that the mode can be used to represent a set of data. She said that it is because the number 22 used to indicate the mode can be counted. She even repeated the similar justification after the number of comments for student D was changed from 22 to 6. Thus, Joan presented that she had the knowledge that the mode can be used as a form of data representation (P) but her justification towards utilising the mode as a form of data representation was found to be inappropriate (I).

In Task 4, Joan mentioned that the average referred to the common number of hours of TV watched per day. She explained that the average is the mean. She elaborated that the average is calculated using total hours of TV watched per day divided with the total number of students. Joan mentioned that the information on the extreme data affects the average.

She said that information on the extreme data can be eliminated from the mean calculation. However, Joan mentioned that in this task, the information would have been included in the mean calculation. She explained that it is because the average is considered for one whole Malaysia, therefore all information should be included in the mean calculation.

In this situation, the average based on the median is the most suitable measure to represent data involving human population. Joan lacked in the knowledge that the average can be based on the median. Thus, she did not present the knowledge of the median as the average (NP). Joan's knowledge of the average is the mean which is acceptable given to proper justification related to the information on extreme data. However, Joan said that in this situation where the average is found for the whole Malaysia, the mean calculation included all information including the information on the extreme data. This is found to be inappropriate (I).

In Task 5, Joan utilised the idea of median to obtain the weight of the median child. Joan said that the data has to be first arranged in ascending order. She said that when the number of data is nine, the median is the middle number. She got the median for question one as 19. She mentioned that the median after 43 kg was included in the data is still 19. She said that she obtained this by adding the two-middle data and then divided the total with two. Therefore, Joan's knowledge of the idea of the median was found to be present (P) and correct (A).

Joan mentioned that the mean is not adequate to represent the weight of all ten children. She explained that the mean that she calculated was 22.2 in which is different from the weights given in the data. Joan mentioned that the median is more appropriate because it lies between the lightest and the heaviest weight. Joan also said that the mean can be considered if the weight of the tenth child is eliminated and the mean is calculated based on the weights of the remaining nine children. Therefore, Joan's knowledge related to the robustness of the measures of central tendency was found to be present (P) and accurate (A).

In Task 6 and Task 7, Joan could construct data sets based on the given average value and the conditions imposed. In Task 6, Joan placed all the prices for the seven bags of crisps as 27. She explained that she found the average by adding all these values up and then divided by seven in which she got back the average as 27. Therefore, Joan's knowledge of constructing a data set based on the idea of average was present (P) and found to be appropriate because she had fallen back on the mean algorithm (A).

In Task 7, Joan placed the values randomly. She said that she added the values that she placed and divided the total with 7 to get the average as 27. She also utilised the backward mean calculation where the average 27, she multiplied with 7 and got the total price as 189 in which she based the prices within the total. Joan's choices for the prices followed the normal distribution bell shape which had prices that started from smaller to bigger values. Thus, Joan's knowledge of constructing a data set based on the idea of average for Task 7 was found to be present (P) and appropriate (A).

Table 4.150

Coding rubrics for determining level of Joan's subject matter knowledge of measures of central tendency with reference to context

Task	Subject Matter Knowledge of Measures of Central Tendency with reference to Context																															
	Mean as average				Mode as average				Median as average				Idea of mode				Idea of median				Mode as data representation				Robustness of measures				Idea of average			
	P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP	
	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I
Task 1	X																															
Task 2								X					X								X											
Task 3																X						X										
Task 4												X																				
Task 5																	X								X							
Task 6																													X			
Task 7																													X			
	1												**1**				**1**				**1**				**1**				**2**			

Legend: P = Present NP = Not Present A = Appropriate I = Inappropriate

Joan's Percentage of Subject Matter Knowledge of Measures of Central Tendency with Reference to Context

$$= \frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)} \text{ X } 100\% = \frac{\mathbf{7}}{\mathbf{11}} \times \mathbf{100\%} = 63.67\% \text{ (Medium)}$$

Level of subject matter knowledge of measures of central tendency in handling bias. In Task 8, Joan said that the mean is calculated by adding all the values and then divide the total with the number of data. However, she said that to get a more accurate mean, she eliminated the extreme value 15.3 and calculated the mean based on the remaining eight data. She got the mean as 6.17. Therefore, Joan not only identified that the mean can be obtained but also summarised the data using the mean by calculating the mean value. Therefore, Joan's knowledge of identifying and summarising the data using the mean was found to be present (P) and correct (A).

Joan said that the median can be found after the data is arranged in ascending order. She said the median was 6.2 which was the middle number in the arrangement. Therefore, Joan's knowledge of identifying and summarising the data using the median was found to be present (P) and correct (A).

Joan also identified the mode as 6.3. She identified the mode based on the highest frequency where 6.3 had the highest frequency. Therefore, Joan's knowledge of identifying and summarising the data using the mode was found to be present (P) and correct (A).

Joan mentioned that the mean is the best measure to represent the actual weight of this object. According to her, in the mean calculation any unnecessary data can be eliminated for better accuracy. Therefore, Joan's knowledge related to the best representation measure is found to be present (P) and her justification towards the best measure was appropriate (A).

Similarly, in Task 9, Joan calculated the mean after eliminating the extreme data 180 000. She added the remaining four data and then divided the total with four. She got the mean as 51 000. Therefore, Joan not only identified that the mean can be obtained but also summarised the data using the mean by calculating the mean value.

Therefore, Joan's knowledge of identifying and summarising the data using the mean was found to be present (P) and correct (A).

Joan said that the median can be found after the data was arranged in ascending order. She said the median was 54 000 which was the middle number in the arrangement. Therefore, Joan's knowledge of identifying and summarising the data using the median was found to be present (P) and correct (A).

Joan mentioned that in this task there is no mode. She explained that the frequency for each salary was one. Therefore, Joan's knowledge of identifying and summarising the data using the mode was found to be present (P) and correct (A).

Joan mentioned that the mean is the best measure to represent the data because the mean excluded the extreme data. Therefore, Joan's knowledge related to the best representation measure is found to be present (P) and appropriate (A).

Table 4.151

Coding rubrics for determining level of Joan's subject matter knowledge of measures of central tendency in handling bias

Task	Subject Matter Knowledge of Measures of Central Tendency in Handling Bias															
	Identifies and summarises data using mean				Identifies and summarises data using median				Identifies and summarises data using mode				Best data representation measure			
	P		NP		P		NP		P		NP		P		NP	
	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I
8	X				X				X				X			
9	X				X				X				X			
	2				**2**				**2**				**2**			

Legend: P = Present NP = Not Present A = Appropriate I = Inappropriate

Joan's Percentage of Subject Matter Knowledge of Measures of Central Tendency in Handling Bias

$= \frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)}$ X 100% = $\frac{\mathbf{8}}{\mathbf{8}} \times \mathbf{100}\%$ = 100% (High)

Level of subject matter knowledge of measures of central tendency in problem solving. In Task 10, Joan carried out the backward mean calculation to obtain the total grade point average for College A based on two semesters as 3.2 times 2. She also found the total grade point average for College B based on three semesters as 3.8 times 3. Therefore, Joan's knowledge of performing the backward mean calculation to obtain the total from the mean is found to present (P) and correct (A).

Joan explained that the mean which was given for each college was for the number of semesters involved in that particular college. She also added the total grade points of both colleges and the total semesters of both colleges before she proceeded in the weighted mean calculation. Thus, Joan's knowledge related to the representative nature of the mean is found to be present (P) and appropriate (A).

Finally, Joan carried out the forward mean calculation to find the weighted mean. She divided the total grade points of both colleges with the total semesters of both colleges and found the weighted mean as 3.56. Therefore, Joan's knowledge of the forward mean calculation is found to present (P) and appropriate (A) despite the task not explicitly stating to find for the weighted mean.

In Task 11, Joan performed the backward mean calculation and obtained the total hours of TV watched per weekend for both rural and urban groups. She said that she multiplied the number of students in each group with the group average. Joan found the total hours of TV watched for rural group by multiplying 25 with 8. Similarly, for the urban group she multiplied 75 with 4. Thus, Joan's knowledge of performing backward mean calculation was found to be present (P) and appropriate (A).

Joan explained that she multiplied the average of each group with the number of students in each group because the given average was for the group. She also mentioned that the totals of both the groups were added and the total number of students of both groups were found before she proceeded with further calculation because she wanted to find the overall average for all the students. Thus, Joan's knowledge related to the representative nature of the mean is found to be present (P) and appropriate (A).

Finally, Joan carried out the forward mean calculation to find the weighted mean. She divided the total number of hours of both groups with the total number of students and found the average as 5. Therefore, Joan's knowledge of the forward mean calculation is found to present (P) and appropriate (A) despite the task not explicitly stating to find for the weighted mean.

Table 4.152

Coding rubrics for determining level of Joan's subject matter knowledge of measures of central tendency in problem solving

Task	Subject Matter Knowledge of Measures of Central Tendency in Problem Solving											
	Backward mean calculation				Representative nature of the mean				Forward mean calculation			
	P		NP		P		NP		P		NP	
	A	I	A	I	A	I	A	I	A	I	A	I
10	X				X				X			
11	X				X				X			
	2				**2**				**2**			

Legend: P = Present NP = Not Present A = Appropriate I = Inappropriate

Joan's Percentage of Subject Matter Knowledge of Measures of Central Tendency in Problem Solving

$$= \frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)} \text{ X } 100\% = \frac{\mathbf{6}}{\mathbf{6}} \times \mathbf{100}\% = 100\% \text{ (High)}$$

Level of subject matter knowledge of measures of central tendency in making inference. In Task 12, Joan calculated the mean scores for both the players by adding all the scores and divided the total with the number of games that they played. She arranged the data in ascending order and then found the median as the average of the two middle numbers in the arrangement.

Joan found that the mean for Ramu was 19.7 and the mean for Beng was 21.8. She also found the median for Ramu as 20.5 and for Beng as 21.5. Therefore, Joan's knowledge of summarising equal sized numerical data using the measures of central tendency without the task explicitly stating to do so was found to be present (P) and appropriate (A).

Joan utilised the mean to decide on the player to be recommended for the team. She said that the mean or the average takes count the overall data which can be a predictor for the next score. Thus, Joan's knowledge of utilising the appropriate measure of central tendency to compare and draw conclusion on two equal sized numerical data was found to be present (P) and appropriate given to her justification (A)

In Task 13, Joan summarised the data using mean and median. She divided the total scores for Ali with eight because he played eight games and the total scores for Khan, she divided with six because he played six games. Therefore, Joan's knowledge of summarising unequal sized numerical data using the measures of central tendency without the task explicitly stating to do so was found to be present (P) and correct (A).

Initially Joan compared the mean and the median to decide on the player to be selected. However, after she realised that the two players played unequal sets of games, Joan decided to use the mean because she said that it is a fair measure and considers the overall games played by the players. Thus, Joan's knowledge of utilising the appropriate measure of central tendency to compare and draw conclusion on two unequal sized numerical data in this case the mean was found to be present (P) and appropriate given to her justification (A)

In Task 14, Joan used the mean and the mode to summarise the data. She calculated the mean by counting the number of students and then divided the total with 4 which she referred to as the number of questions. She got the mean for both the classes as 2.25. Joan found the mode for 5B as 4 and mentioned that 5A had no mode.

Joan utilised the mean and the mode to summarise two equal sized graphical data without the task explicitly stating to do so. Therefore, Joan's knowledge of summarising two equal sized graphical data using the measures of central tendency was found to be present (P). However, Joan calculated the mean for both the classes incorrectly and also identified the mode for 5A incorrectly. Thus, although she presented the knowledge of utilising the measures of central tendency as a data summary tool but she had incorrectly calculated mean and incorrectly identified the mode for 5A (I).

Joan had utilised the mode to compare and decide on the player to be selected for the team. She selected 5B because she claimed that 5B had a mode of 4 whereas 5A had no mode. Joan did not utilise the mean because she found that there was no difference in the mean values for both the classes. The mean is the best choice to compare and draw conclusion based on the performance of groups. However, Joan had calculated the mean incorrectly which made her to decide not to use the mean. Therefore, Joan's knowledge of utilising the appropriate measure to compare and draw conclusion on two equal sized graphical data was found to be not present (NP) and problematic (I)

In Task 15, Joan utilised the mean to summarise two unequal sized graphical data. She calculated the mean by adding the number of students and then she divided the total with 8 which she referred to as the number of questions. She found that the mean for the class 5C as 4.5 and the mean for the class 5D as 2.6. Therefore, Joan's knowledge of summarising two unequal sized graphical data using the measures of central tendency in this case the mean was marked to be present (P). However, Joan had calculated the mean for both the classes incorrectly (I).

Joan compared the mean for both the classes and the class that had the bigger mean was chosen as the class that performed better. She found that 5C performed better because of the bigger average. Joan said that she preferred to use the mean because the classes had unequal number of students. Therefore, Joan's knowledge of utilising the appropriate measure of central tendency in this case the mean to compare and draw conclusion about two unequal sized graphical data was found to be present (P). Joan's justification towards the use of the mean as the appropriate measure to compare and draw conclusions based on two unequal sized graphical data was appropriate too (A).

Table 4.153

Coding rubrics for determining level of Joan's subject matter knowledge of measures of central tendency in making inference

Task	Subject Matter Knowledge of Measures of Central Tendency in Making Inference																															
	Summarises two equal sized numerical data using measures of central tendency				Utilises the appropriate measure to compare and draw conclusion about two equal sized numerical data				Summarises two unequal sized numerical data using measures of central tendency				Utilises the appropriate measure to compare and draw conclusion about two unequal sized numerical data				Summarises two equal sized graphical data using measures of central tendency				Utilises the appropriate measure to compare and draw conclusion about two equal sized graphical data				Summarises two unequal sized graphical data using measures of central tendency				Utilises the appropriate measure to compare and draw conclusion about two unequal sized graphical data			
	P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP	
	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I
Task 12	X				X																											
Task 13									X				X																			
Task 14																		X						X								
Task 15																										X			X			
	1				**1**				**1**				**1**																**1**			

Legend: P = Present NP = Not Present A = Appropriate I = Inappropriate

Joan's Percentage of Subject Matter Knowledge of Measures of Central Tendency in Making Inference

$= \frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)}$ X 100% = $\frac{\mathbf{5}}{\mathbf{8}} \times \mathbf{100\%}$ = 62.5% (Medium)

Joan's mean percentage of subject matter knowledge for all four constructs $\frac{\mathbf{63.67+100+100+62.5}}{\mathbf{4}} = \mathbf{81.54\%}$. Therefore, Joan's level of subject matter knowledge of measures of central tendency is medium

Chapter 5 Summary of Findings

Introduction

The purpose of this study was to investigate pre-service mathematics teachers' subject matter knowledge (SMK) of measures of central tendency. The present study attempted to answer the following research questions:

1. What types of subject matter knowledge of measures of central tendency do the pre-service mathematics teachers' have?

2. What levels of subject matter knowledge of measures of central tendency do the pre -service mathematics teachers' exhibit?

This chapter provides the summary of the findings across cases under five themes namely: average, data representation, procedures, data summary, and data comparison.

Summary of the Findings

The findings in this study were based on the data collected from the audio and video recordings during the clinical interviews. The participants' notes in the task sheets were used to further validate the findings. In this section, to answer research question one, findings of pre-service mathematics teachers' subject matter knowledge of measures of central tendency were summarised in terms of the identified themes that were emerged from the clinical interview.

To answer research question two, findings of pre-service mathematics teachers' levels of subject matter knowledge of measures of central tendency were summarised in terms of its level for each of the four constructs of investigation namely; with reference to the context, in handling bias, in problem solving, and in making inference. Based on this, the levels of subject matter knowledge of measures of central tendency in terms of the identified themes that were emerged from the clinical interview were summarised. The overall level of subject matter knowledge of measures of central tendency both in relation to the four constructs of investigation and according to the identified themes from the clinical interview were also calculated.

Measures of central tendency as averages. In this subsection, the findings of pre-service mathematics teachers' subject matter knowledge of measures of central tendency were summarised under the theme measures of central tendency as averages based on the following categories: (a) mean as average, (b) mode as average, (c) median as average, and (d) idea of average. Table 5.1 shows a summary of pre-service mathematics teachers' subject matter knowledge of measures of central tendency under the theme measures of central tendency as averages.

Mean as average. Finding of this study suggests that all six pre-service mathematics teachers held the knowledge of the mean as average. This was clearly seen in their responses where all six of them had mentioned "sum and division" or "add and divide" when they were asked on how the average would have been obtained. In fact, two of them, Amy and Tina had clearly mentioned that the average is the mean.

The pre-service mathematics teachers also showed mean calculation in their examples. Furthermore, all six pre-service mathematics teachers mentioned that the "use of mean algorithm", "the division of any two numbers", "the division of two non-divisible numbers", and "the division of even odd numbers" involved in the mean calculation as the reason behind a decimal average.

Mode as average. Only one out of the six pre-service mathematics teachers agreed to utilise the mode as the average. He said that since the mode indicated the most, therefore it can be utilised as the average. The rest disagreed on utilising the mode as the average. They mentioned that the average should be based on the mean. Whereas, one of them mentioned that the mode can be used as a guide to decide on the suitable average value.

Median as average. In relation to the use of the median in the reporting of an average involving data related to the human population, none of the pre-service mathematics teachers had mentioned that the average can be based on the median. Two out of the six mentioned that the average can be based on the mode because of the existence of extreme data.

Meanwhile the remaining four mentioned that the average was obtained from the mean calculation in which all the data including the extreme data were considered in the calculation. These pre-service mathematics teachers also believed that if the average is referring to a population such as Malaysia then all the data has to be taken into account in the mean calculation.

Idea of average. Regarding the pre-service mathematics teachers' idea of the average, it can be said that all them had algorithmic relationship to the idea, both in the restriction and none restriction cases. These teachers had utilised the mean algorithm to calculate the total from the given average and utilised the total to guide their construction.

However, it was found that for the case where no restriction was posed on using the average value back in the construction, five out of the six pre-service mathematics teachers had used the average price for all the prices in the construction which indicated the idea of average as modal.

One of them had constructed based on the values which are reasonably close to the given average and another one had showed the idea of the average as a balance point where she balanced up pairs of prices to get the average.

Whereas, for the case where a restriction on utilising back the average price in the construction was posed, three out of the six pre-service mathematics teachers had exhibited the idea of average as reasonable sense. Two had only utilised the mean algorithm to guide their construction with no sense towards the prices considered. And again, only one of the pre-service mathematics teacher had showed the idea of average as a balance point where she balanced up the prices in pairs to get the average price. She also exhibited the similar idea in the earlier none restriction case.

Table 5.1

Summary of pre-service mathematics teachers' subject matter knowledge of measures of central tendency as averages

Category	Findings	Evidence	Subject
Mean as average	Mean calculation	Mentioned "sum and division", "add and divide".	Lina, Bella, Harry, Amy, Tina, Joan,
	Mean is average	Mentioned that the mean is the average	Amy, Tina
		Clear examples that showed the mean calculation	Lina, Bella, Harry, Amy, Tina, Joan
	Decimal average as a result of mean calculation	Mentioned the use of mean algorithm, the division of any two numbers, the division of two non-divisible numbers and the division of even odd numbers in the mean calculation	Lina, Bella, Harry, Amy, Tina, Joan
Mode as average	Mode can be utilised as average	Mentioned that the mode can be utilised as average	Harry
	Mode cannot be utilised as average	Mentioned that the mode cannot be used as the average but can be used to guide	Lina

		on the decision for the suitable average value	
		Totally disagreed on utilising the mode as the average. Instead, mentioned that the average should be based on the mean.	Bella, Amy, Tina, Joan
Median as average	Median can be used in the reporting of an average involving the human population	Mentioned that the average can be based on the mode because the existence of outliers can affect the average	Lina, Harry
		Mentioned that the average was based on the mean calculation that considered all data including the outliers. A belief that when the average is for population such as Malaysia, then all the given data has to be considered in the average.	Bella, Amy, Tina, Joan
Idea of average (none restriction case)	Average as modal with algorithmic relationship	Exhibited that the average can be based on the mode. However, the mean algorithm was still utilised.	Lina, Harry, Amy, Tina, Joan
	Average as algorithmic relationship with reasonable sense	Exhibited that the average was based on the mean algorithm but the values based close to the given average	Bella
	Average as a balance point with algorithmic relationship	Exhibited that the average can be some sort of balance point between the data point but did fall back on the algorithm.	Amy
Idea of average (restriction case)	Average as algorithmic relationship with reasonable sense	Mean algorithm was utilised to guide the construction but the idea of reasonable sense was put forth when the values were determined.	Lina, Bella, Tina

	Average as algorithmic relationship	Mean algorithm was utilised to guide the construction	Harry, Joan
	Average as algorithmic relationship with balance point	Adjusted the points from the earlier construction based on the average value.	Amy

Measures of central tendency as data representation. In this subsection, the findings of pre-service mathematics teachers' subject matter knowledge of measures of central tendency were summarised under the theme measures of central tendency as data representation based on the following categories: (a) mode as data representation, (b) best data representation measure, and (c) robustness of measures. Table 5.2 shows a summary of pre-service mathematics teachers' subject matter knowledge of measures of central tendency as data representation.

Mode as data representation. Four out of six pre-service mathematics teachers in this study had some knowledge that the mode can be utilised as a form of data representation. However, their justification towards its use varied. Lina mentioned that she would use the mode as a form of data representation because it indicated the highest demand or favored by many. However, she also mentioned that its use is subjected to it being close to the other data too. Whereas, Bella and Amy mentioned that they would only use the mode as a form of data representation if the mode and the average are around the same value. Meanwhile Joan mentioned that the mode can be utilised because it can be counted.

On the other hand, Harry and Tina totally disagreed that the mode can be a form of data representation because the mode was based on one particular data with the highest frequency. According to them, the suitable form of data representation has to consider all the data.

Best data representation measure. Two out of the six pre-service mathematics teachers mentioned that the median is the best representation measure in the case where there existed extreme data. However, these pre-service mathematics teachers also mentioned that the mean can be considered if the extreme data was excluded from the mean calculation.

Whereas Lina, mentioned that either the median or the mode is suitable if these measures are representative of the data. She totally disagreed on using the mean to represent the data in the presence of an extreme data. However, she did not mention if the mean which was calculated after excluding the extreme data can be considered.

On the other hand, two of the pre-service mathematics teachers, Tina and Joan mentioned that the mean that excluded the extreme data is the best measure to represent the data because it considered all the data. Tina mentioned that this was unlike for the median which considered only the middle data and the mode which is based on the data with the highest frequency, thus not suitable to be considered as the best measure to represent the data. Whereas Joan, believed that the median and the mode which were based on the data that still included the extreme data should not be considered as the best measure to represent the data.

Bella on the other hand also believed that the best representation measure has to consider all the given data and the mean is the most suitable measure because it considered all the data. However, she mentioned that the extreme data has to be tested first to see if it needed to be included or excluded from the mean calculation.

Robustness of measures. Four out of six pre-service mathematics teachers had the knowledge that the median is a more robust measure in representing the data than the mean in the presence of an extreme data. They were aware that an extreme data can affect the mean and that the median is more suitable to represent the data in such cases. However, these teachers also agreed that if the extreme data was discarded from the mean calculation than the mean is adequate to represent the data.

Only one pre-service mathematics teacher did not mention on the median being a more adequate measure than the mean in the presence of an extreme data. Instead she said that the mean which is calculated after excluding the extreme data is adequate to represent the data. On contrary, another per-service mathematics teacher mentioned that the mean that considered all the data including the extreme is adequate to represent the data.

Table 5.2

Summary of pre-service mathematics teachers' subject matter knowledge of measures of central tendency as data representation

Category	Findings	Evidence	Subject
Mode as data representation	Mode can be used as a form of data representation.	Mentioned that because "majority" bought that size or the mode indicated highest demand or favored by many	Lina
		However, the mode can be utilised only if the mode is close to the other data points Mentioned that can be utilised when the mode and the average are around the same.	Bella, Amy
		Mode can be counted	Joan
	Mode cannot be utilised as a form of data representation	Mode was based on one particular data with the highest frequency. The other data points need to be considered too in data representation.	Harry, Tina

Best data representation measure	Best representation measure was based on the measures that the values were around the values given in the data.	Mentioned that in the presence on an outlier, the mean that considered all the data is not representative of the data. Instead mentioned that the median or the mode is more suitable if these measures are representative of the data.	Lina
	Best representation measure had to consider all the data including the extreme.	Mentioned that the best measure had to consider all the given data and the mean is the most suitable. However, the subject also mentioned that the extreme value must be tested to see if it can be rejected from the mean calculation to obtain the best measure	Bella
	Best representation measure is the median in the case of the presence of an extreme data	Mentioned that the median is the best measure to represent a set of data that contained extreme data. However, the subject mentioned that the mean can also be considered if the extreme is excluded from the mean calculation.	Harry, Amy
	Best representation measure is the mean that was calculated after excluding the extreme data	Mentioned that the mean that excluded the extreme data is the best measure to represent the data because it considered all the other data unlike the median which considered the middle data and the mode based on the data with the highest frequency	Tina
		Mentioned that the mean that excluded the extreme data is the best measure to represent the data	Joan

		because it considered all the other data unlike the median or the mode which was based on the data which still included the extreme data.	
Robustness of measures	Median is adequate to represent the data than the mean in the presence of an extreme data	Mentioned that in the presence of an extreme data the mean is not adequate to represent the data. Instead the median is more suitable. Mean can be considered if the extreme data is discarded first in the mean calculation.	Lina, Harry, Amy, Joan
	Mean considered adequate after excluding the extreme data	Mentioned that the mean calculated after excluding the extreme data is adequate to represent the data	Bella
	Mean which considered all the data including the extreme data	Mentioned that the mean that considered all the data including the extreme data is adequate to represent the data	Tina

Measures of central tendency procedures. In this subsection, the findings of pre-service mathematics teachers' subject matter knowledge of measures of central tendency were summarised under the theme measures of central tendency procedures based on the following categories: (a) idea of mode, (b) idea of median, (c) backward mean calculation, representative nature of the mean, and (d) forward mean calculation. Table 5.3 shows a summary of pre-service mathematics teachers' subject matter knowledge of measures of central tendency procedures.

Idea of mode. All six pre-service mathematics teachers related the idea of mode in real life situations without the tasks explicitly stating to do so. They utilised the frequency tally method to find the frequency of each data in which later they identified the mode based on the highest frequency.

However, when the per-service mathematics teachers were presented with a situation on the number of comments made by several students, only Lina identified the mode correctly as 2 based on the number of comments which had the highest frequency or "the most frequent number of comments".

Two of the pre-service mathematics teachers identified the mode based on the biggest number of comments given in the table in which they believed that this number indicated the highest frequency for the number of comments. Whereas, the remaining three identified the mode as student D in which they based on the category that carried the highest frequency. However, these pre-service teachers also believed that the biggest number of comments indicated the highest frequency for the number of comments.

Idea of median. Five out of six pre-service mathematics teachers carried out the correct median procedure to find the median both for odd and even numbered data. However, out of these five teachers, only one mentioned that the data can be arranged in both ascending and descending order before the median is found. The rest four mentioned that the data should be arranged in ascending order.

On the other hand, one of the pre-service mathematics teacher had lacked in the knowledge of the median procedure. She elaborated incorrectly the procedure to find the median in which the median that she found was also incorrect.

Backward mean calculation. All six pre-service mathematics teachers utilised the backward mean calculation to find the total for the groups from the given group averages.

Representative nature of the mean. All six pre-service mathematics teachers have exhibited the knowledge that the data sets needed to be combined in order to get the weighted mean.

Forward mean calculation. All six pre-service mathematics teachers have carried out the forward mean calculation to get the weighted mean without the tasks explicitly stating to do so.

Table 5.3

Summary of pre-service mathematics teachers' subject matter knowledge of measures of central tendency procedures

Category	Findings	Evidence	Subject
Idea of mode	Related the idea of mode in real life situation.	Frequency tally used to find the frequency of each shoe size and identified the mode based on the highest number of purchase.	Lina, Bella, Harry, Amy, Tina, Joan
	Mode 2 based on the "most frequent"	Most frequent was identified based on the highest frequency. 2 had the highest frequency where it occurred thrice compared to the other	Lina

		data that occurred only once.	
	Mode 22 based on the "biggest number"	"Biggest number" was believed to indicate the highest frequency. In Task 3, the biggest number of comments 22 was identified as the mode.	Bella, Tina
	Mode as student D based on the category that referred to the "biggest number"	The category that carried the "biggest number"	Harry, Amy, Joan
Idea of median	Correct median procedure was carried out both for even and odd numbered data	Mentioned that the data is rearranged in ascending or descending order first. Then if the data is odd numbered, the median is the middle number whereas if the data is even numbered, the median is the found by adding the two middle numbers and then divided by two.	Lina
		Mentioned that the data is rearranged in only ascending order first. Then if the data is odd numbered, the median is the middle number whereas if the data is even numbered, the median is the found by adding the two middle numbers and then divided by two.	Harry, Amy, Tina, Joan
	Incorrect median procedure was carried out both for even and odd numbered data	Mentioned that the data was arranged in ascending order first. Then added all the values and divided the total with two. This value obtained was checked for its place position in the ordered arrangement by adding all weights in the rearrangement one by one till the position of this value	Bella

		is obtained. Got the median by adding the two values in between the position.	
Backward mean calculation	The total for the groups from the given group averages.	Obtained the total for the groups from the given average without the task explicitly stating to do so.	Lina, Bella, Harry, Amy, Tina, Joan
Representative nature of the mean	Representative nature of the mean in making sense the need for the data sets to be combined in weighted mean.	Combined the group totals and exhibited that the data sets needed to be combined in weighted mean	Lina, Bella, Harry, Amy, Tina, Joan
Forward mean calculation	Calculated the weighted mean	Carried out the forward mean calculation using the combined totals and obtained the weighted mean	Lina, Bella, Harry, Amy, Tina, Joan

Measures of central tendency as data summary. In this subsection, the findings of pre-service mathematics teachers' subject matter knowledge of measures of central tendency were summarised under the theme measures of central tendency as data summary based on the following categories: (a) identifies and summarises data using mean, (b) identifies and summarises data using median, (c) identifies and summarises data using mode, (d) summarises equal sized numerical data using measures of central tendency, (e) summarises unequal sized numerical data using measures of central tendency, (f) summarises equal sized graphical data using measures of central tendency, and (g) summarises unequal sized graphical data using measures of central tendency. Table 5.4 shows a summary of pre-service mathematics teachers' subject matter knowledge of measures of central tendency as data summary.

Identifies and summarises data using mean. All six pre-service mathematics teachers had identified that the mean can be used to summarise the given data. All of them had utilised the mean algorithm correctly to calculate the mean.

Identifies and summarises data using median. All six pre-service mathematics teachers had identified that the median can be obtained from the given data. However, only five out of the six were able to carry out the median procedure correctly to obtain the median. Bella carried out the median procedure incorrectly and obtained an incorrect median.

Identifies and summarises data using mode. Five out of six pre-service mathematics teachers had identified the mode based on the highest frequency. They also mentioned that for the data where all the data occurred once, no mode existed. These teachers were able to justify correctly why there was no mode in such situation which they explained was because each data occurred once and there was no data with highest frequency. However, one of the pre-service mathematics teacher, Bella had two different ways of identifying the mode. She identified the mode based on the highest frequency and she also identified the mode based on the biggest value that existed in the data for two different sets of data which were given to her in the task 8 and task 9.

Summarises equal sized numerical data using measures of central tendency. Four out of six pre-service mathematics teachers summarised the given equal sized numerical data using the mean only. Whereas one pre-service mathematics teacher summarised using the mean and the median. Only one pre-service mathematics teacher summarised the given data using all three measures of central tendency.

Summarises unequal sized numerical data using measures of central tendency. Similar to the equal sized data, four out of six pre-service mathematics teachers summarised the given unequal sized numerical data using the mean only. Whereas one pre-service mathematics teacher summarised using the mean and the median and another one pre-service mathematics teacher summarised the given data using all three measures of central tendency.

Summarises equal sized graphical data using measures of central tendency. Three out of six pre-service mathematics teachers summarised equal sized graphical data using the mean. Two of the pre-service mathematics teachers summarised the data using the mode and only one of the pre-service mathematics teachers summarised the data using all three measures of central tendency.

Summarises unequal sized graphical data using measures of central tendency. Three out of the six pre-service mathematics teachers summarised the given unequal sized graphical data using the mean. One of them summarised the data using the mean and the mode. Whereas one summarised using only the mode. However, one of the pre-service mathematics teachers did not summarise the given data using any of the three measures of central tendency.

Table 5.4

Summary of pre-service mathematics teachers' subject matter knowledge of measures of central tendency as data summary

Category	Findings	Evidence	Subject
Identifies and summarises data using mean	Mean algorithm	Identified and utilised the mean algorithm to calculate the mean	Lina, Bella, Harry. Amy. Tina, Joan
Identifies and summarises data using median	Median procedure	Identified and carried out the median procedure for odd numbered data appropriately.	Lina, Harry, Amy, Tina, Joan
		Identified the median but carried out the wrong procedure to find the median	Bella
Identifies and summarises data using mode	Mode identification	Identified the mode based on the highest frequency	Lina, Harry, Amy, Tina, Joan
		Identified the mode based on the highest frequency and also the biggest value.	Bella
Summarises equal sized numerical data	Utilised all or any one of the measures of central tendency to	Summarised the given data using all three	Lina

using measures of central tendency	summarise equal sized numerical data without the task explicitly stating to do so	measures of central tendency	
		Summarised the given data using mean	Bella, Harry, Amy, Tina
		Summarised the given data using mean and median	Joan
Summarises unequal sized numerical data using measures of central tendency	Utilised all or any one of the measures of central tendency to summarise unequal sized numerical data without the task explicitly stating to do so	Summarised the given data using all three measures of central tendency	Lina
		Summarised the given data using mean	Bella, Harry, Amy, Tina
		Summarised the given data using mean and median	Joan
Summarises equal sized graphical data using measures of central tendency	Utilised all or any one of the measures of central tendency to summarise equal sized graphical data without the task explicitly stating to do so	Summarised the given data using all three measures of central tendency	Lina
		Summarised the given data using mean	Bella, Amy, Tina
		Summarised the given data using mode	Harry, Joan
Summarises unequal sized graphical data using measures of central tendency	Utilised all or any one of the measures of central tendency to summarise unequal sized graphical data without the task explicitly stating to do so	Summarised the given data using the mode	Lina
		Summarised the given data using the mean	Bella, Tina, Joan
		Summarised the given data using mean and mode	Amy
		Did not summarise the given data using any of the three measures	Harry

Measures of central tendency as data comparison. In this subsection, the findings of pre-service mathematics teachers' subject matter knowledge of measures of central tendency were summarised under the theme measures of central tendency as data comparison based on the following categories: (a) utilises the appropriate measure to compare and draw conclusion about equal sized numerical data, (b) utilises the appropriate measure to compare and draw conclusion about unequal sized numerical data, (c) utilises the appropriate measure to compare and draw conclusion about equal sized graphical data, and (d) utilises the appropriate measure to compare and draw conclusion about unequal sized graphical data. Table 5.5 shows a summary of pre-service mathematics teachers' subject matter knowledge of measures of central tendency as data comparison.

Utilises the appropriate measure to compare and draw conclusion about equal sized numerical data. Five out of the six pre-service mathematics teachers utilised the mean to compare and draw conclusion based on the two-equal sized numerical data. They even justified that the mean is the best measure to draw conclusion. Only one of the pre-service mathematics teachers did not utilise the mean. Instead, she compared the mode to draw the conclusion based on the data sets.

Utilises the appropriate measure to compare and draw conclusion about unequal sized numerical data. Similar to the equal sized numerical data, five out of the six pre-service mathematics teachers utilised the mean to compare and draw conclusion based on the two-unequal sized numerical data. They even justified that the mean is the best measure to draw conclusion. Only one of the pre-service mathematics teachers did not utilise the mean. Instead, she compared the mode to draw the conclusion based on the data sets.

Utilises the appropriate measure to compare and draw conclusion about equal sized graphical data. Three out of the six pre-service mathematics teachers utilised the mean to compare and draw conclusions based on two equal sized graphical data. Whereas, the other three compared the mode to draw conclusion.

Utilises the appropriate measure to compare and draw conclusion about unequal sized graphical data. Four out of the six pre-service mathematics teachers utilised the mean to compare and draw conclusion based on two unequal sized graphical data. One of them utilised the mode. Whereas one did not utilise any of the three measures to compare and draw conclusion based on the given data sets.

Table 5.5

Summary of pre-service mathematics teachers' subject matter knowledge of measures of central tendency as data comparison

Category	Findings	Evidence	Subject
Utilises the appropriate measure to compare and draw conclusion about equal sized numerical data	Utilised the appropriate measure of central tendency to compare and draw conclusion.	Chose mode comparison	Lina
		Chose mean comparison and explained as the best to draw conclusion	Bella, Harry, Amy, Tina, Joan
Utilises the appropriate measure to compare and draw conclusion about unequal sized numerical data	Utilised the appropriate measure of central tendency to compare and draw conclusion.	Chose mode comparison	Lina
		Chose mean comparison and explained as the best to draw conclusion	Bella, Harry, Amy, Tina, Joan
Utilises the appropriate measure to compare and draw conclusion about equal sized graphical data	Utilised the appropriate measure of central tendency to compare and draw conclusion.	Chose mode comparison	Lina, Harry, Joan
		Chose mean comparison and explained as the best to draw conclusion	Bella, Amy, Tina
Utilises the appropriate measure to compare and draw conclusion about unequal sized graphical data	Utilised the appropriate measure of central tendency to compare and draw conclusion.	Chose the mode comparison	Lina
		Chose the mean comparison	Bella, Amy, Tina, Joan

	Did not utilise any of the three measures to compare and draw conclusion	Harry

Level of subject matter knowledge. With regard to the subject matter knowledge of measures of central tendency in terms of the four constructs of investigation, all six pre-service mathematics teachers attained a medium level for the construct with reference to context. In relation to the subject matter knowledge of measures of central tendency in handling bias, four of the pre-service mathematics teachers attained a high level and the remaining two attained a medium level. For the subject matter knowledge of measures of central tendency in problem solving, all six pre-service mathematics teachers attained a high level. Meanwhile for the subject matter knowledge of measures of central tendency in making inference, two of the pre-service mathematics teachers attained a high level and the remaining four attained a medium level.

In relation to the overall level of subject matter knowledge of measures of central tendency in terms of the four constructs of investigation, it was found that four out of six pre-service mathematics teachers secured an overall high level of subject matter knowledge of measures of central tendency while the remaining two secured an overall medium level. Table 5.6 illustrates a summary of the pre-service mathematics teachers' levels of subject matter knowledge namely, with reference to the context, in handling bias, in problem solving, and in making inference as well as the overall level of subject matter knowledge of measures of central tendency in terms of the four constructs of investigation.

Based on these levels of subject matter knowledge of measures of central tendency in terms of the four constructs of investigation, the levels of the subject matter knowledge of measures of central tendency in terms of the identified themes were constructed. For the level of subject matter knowledge of measures of central tendency as averages, two of the pre-service mathematics teachers attained a high level whereas the remaining four attained a medium level as shown in Table 5.7. For data representation, two of the pre-service mathematics teachers attained a high level, three attained a medium level and one attained a low level as illustrated in Table 5.8.

With regard to the subject matter knowledge of measures of central tendency procedures five out of the six pre-service mathematics teachers attained a high level and the remaining one attained a medium level as shown in Table 5.9. For data summary, five out of six pre-service mathematics teachers attained a high level whereas the remaining one attained a medium level as shown in Table 5.10. Finally, for the subject matter knowledge of measures of central tendency as data comparison, three pre-service mathematics teachers attained a high level, two attained a medium level and one attained a low level as illustrated in Table 5.11.

However, when the cross cases comparison was done in terms of the emerged themes, only one out of the six pre-service mathematics teachers attained an overall high level. The remaining five attained an overall medium level. Table 5.12 illustrates a summary of the pre-service mathematics teachers' levels of subject matter knowledge of measures of central tendency for each of the emerged themes namely; averages, data representation, procedures, data summary and data comparison as well as the overall level of subject matter knowledge of measures of central tendency in terms of the emerged themes.

Table 5.6

Summary of pre-service mathematics teachers' levels of subject matter knowledge of measures of central tendency with reference to context, in handling Bias, in problem solving and in making Inference as well as the overall level of subject matter knowledge of measures of central tendency

Level of SMK with Reference to Context			Level of SMK in Handling Bias			Level of SMK in Problem Solving			Level of SMK in Making Inference			Overall Level of SMK			Subject
L	M	H	L	M	H	L	M	H	L	M	H	L	M	H	
	X				X			X		X				X	Lina
	X			X				X			X		X		Bella
	X			X				X		X			X		Harry
	X				X			X			X			X	Amy
	X				X			X		X				X	Tina
	X				X			X		X				X	Joan

Legend: L = Low M = Medium H = High

Table 5.7

Summary of pre-service mathematics teachers' levels of subject matter knowledge of measures of central tendency as averages

Subject	Subject Matter Knowledge of Measures of Central Tendency as Average																Level of SMK of measures of central tendency as averages $\frac{n(PA+NPA)}{n(PA+PI+NPA+NPI)} \times 100\%$
	Mean as average				Mode as average				Median as average				Idea of average				
	P		NP		P		NP		P		NP		P		NP		
	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	
Lina	X(T1)					X(T2)					X(T4)		X(T5) X(T6)				80%
Bella	X(T1)							X(T2)				X(T4)	X(T5) X(T6)				60%
Harry	X(T1)				X(T2)						X(T4)		X(T5) X(T6)				100%
Amy	X(T1)							X(T2)				X(T4)	X(T5) X(T6)				60%
Tina	X(T1)							X(T2)				X(T4)	X(T5) X(T6)				60%
Joan	X(T1)							X(T2)				X(T4)	X(T5) X(T6)				60%

Table 5.8

Summary of pre-service mathematics teachers' levels of subject matter knowledge of measures of central tendency as data representation

Subject	Subject Matter Knowledge of Measures of Central Tendency as Data Representation												
	Mode as data representation				Best data representation				Robustness of measures				Level of SMK of measures of central tendency as data representation $\frac{n(PA+NPA)}{n(PA+PI+NPA+NPI)} \times 100\%$
	P		NP		P		NP		P		NP		
	A	I	A	I	A	I	A	I	A	I	A	I	
Lina	X(T2)	X(T3)			X(T8) X(T9)				X(T5)				80%
Bella	X(T2)			X(T3)	X(T9)	X(T8)					X(T5)		60%
Harry	X(T2)			X(T3)		X(T8) X(T9)				X(T5)			20%
Amy				X(T2) X(T3)	X(T8) X(T9)				X(T5)				60%
Tina				X(T2) X(T3)	X(T8) X(T9)				X(T5)				60%
Joan	X(T2)	X(T3)			X(T8) X(T9)				X(T5)				80%

Table 5.9

Summary of pre-service mathematics teachers' levels of subject matter knowledge of measures of central tendency procedures

Subject	Subject Matter Knowledge of Measures of Central Tendency Procedures																				
	Idea of mode				Idea of median				Backward mean calculation				Representative nature of the mean				Forward mean calculation				Level of SMK of measures of central tendency procedures $\frac{n(PA+NPA)}{n(PA+PI+NPA+NPI)} \times 100\%$
	P		NP		P		NP		P		NP		P		NP		P		NP		
	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	
Lina	X(T2) X(T3)					X(T5)			X(T10) X(T11)				X(T10) X(T11)				X(T10) X(T11)				88.9%
Bella	X(T2)			X(T3)				X(T5)	X(T10) X(T11)				X(T10) X(T11)				X(T10) X(T11)				77.8%
Harry	X(T2)			X(T3)	X(T5)				X(T10) X(T11)				X(T10) X(T11)				X(T10) X(T11)				88.9%
Amy	X(T2)			X(T3)	X(T5)				X(T10) X(T11)				X(T10) X(T11)				X(T10) X(T11)				88.9%
Tina	X(T2)			X(T3)	X(T5)				X(T10) X(T11)				X(T10) X(T11)				X(T10) X(T11)				88.9%
Joan	X(T2)			X(T3)	X(T5)				X(T10) X(T11)				X(T10) X(T11)				X(T10) X(T11)				88.9%

Table 5.10

Summary of pre-service mathematics teachers' levels of subject matter knowledge of measures of central tendency as data summary

Subject	Subject Matter Knowledge of Measures of Central Tendency as Data Summary																												Level of SMK of measures of central tendency as data summary $\frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)} \times 100\%$
	Identifies and summarises data using mean				Identifies and summarises data using median				Identifies and summarises data using mode				Summarises two equal sized numerical data using measures of central tendency				Summarises two unequal sized numerical data using measures of central tendency				Summarises two equal sized graphical data using measures of central tendency				Summarises two unequal sized graphical data using measures of central tendency				
	P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		P		NP		
	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	
Lina	X (T8) X (T9)				X (T8) X (T9)				X (T8) X (T9)				X (T12)				X (T13)				X (T14)				X (T15)				100%
Bella	X (T8) X (T9)							X (T8) X (T9)	X (T8)			X (T9)	X (T12)				X (T13)				X (T14)				X (T15)				70%
Harry	X (T8) X (T9)				X (T9)	X (T8)			X (T8) X (T9)				X (T12)				X (T13)				X (T14)							X (T15)	80%
Amy	X (T8) X (T9)				X (T8) X (T9)				X (T8) X (T9)				X (T12)				X (T13)				X (T14)				X (T15)				100%
Tina	X (T8) X (T9)				X (T8) X (T9)				X (T8) X (T9)				X (T12)				X (T13)							X (T14)				X (T15)	80%
Joan	X (T8) X (T9)				X (T8) X (T9)				X (T8) X (T9)				X (T12)				X (T13)					X (T14)				X (T15)			80%

Table 5.11

Summary of pre-service mathematics teachers' levels of subject matter knowledge of measures of central tendency as data comparison

Subject	Subject Matter Knowledge of Measures of Central Tendency as Data Comparison																
	Utilises the appropriate measure to compare and draw conclusion about two equal sized numerical data				Utilises the appropriate measure to compare and draw conclusion about two unequal sized numerical data				Utilises the appropriate measure to compare and draw conclusion about two equal sized graphical data				Utilises the appropriate measure to compare and draw conclusion about two unequal sized graphical data				Level of SMK of measures of central tendency as data comparison $\frac{n(PA + NPA)}{n(PA + PI + NPA + NPI)} \times 100\%$
	P		NP		P		NP		P		NP		P		NP		
	A	I	A	I	A	I	A	I	A	I	A	I	A	I	A	I	
Lina				X(T12)				X(T13)				X(T14)				X(T15)	0%
Bella	X(T12)				X(T13)				X(T14)				X(T15)				100%
Harry	X(T12)				X(T13)							X(T14)				X(T15)	50%
Amy	X(T12)				X(T13)				X(T14)				X(T15)				100%
Tina	X(T12)				X(T13)				X(T14)				X(T15)				100%
Joan	X(T12)				X(T13)							X(T14)	X(T15)				75%

Table 5.12

Summary of pre-service mathematics teachers' levels of subject matter knowledge of measures of central tendency as averages, data representation, procedures, data summary, and data comparison as well as the overall level of subject matter knowledge of measures of central tendency

Level of SMK of MCT as Averages			Level of SMK of MCT as Data Representation			Level of SMK of MCT as Procedures			Level of SMK of MCT as Data Summary			Level of SMK of MCT as Data Comparison			Overall Level of SMK of MCT			Subject
L	M	H	L	M	H	L	M	H	L	M	H	L	M	H	L	M	H	
		X			X			X			X	X				X		Lina
	X			X			X			X				X		X		Bella
		X	X					X			X		X			X		Harry
	X			X				X			X			X			X	Amy
	X			X				X			X			X		X		Tina
	X				X			X			X		X			X		Joan

Legend: L = Low M = Medium H = High

Chapter 6 Discussion and Conclusion

Introduction

This chapter is organized into four main sections: discussion and conclusions, implications of the findings, recommendations for further research, and contribution of the study.

Discussion and Conclusions

The subject matter knowledge of measures of central tendency from the six pre-service mathematics teachers was summarised in terms of the five main themes that emerged from the findings as listed below. This section includes the discussion and conclusions of the five main themes:

a. Subject matter knowledge of measures of central tendency as averages

b. Subject matter knowledge of measures of central tendency as data representation

c. Subject matter knowledge of measures of central tendency procedures

d. Subject matter knowledge of measures of central tendency as data summary

e. Subject matter knowledge of measures of central tendency as data comparison

In relation to the *subject matter knowledge of measures of central tendency as averages*, the finding of this study shows that all six pre-service mathematics teachers held the knowledge of the *mean as the average*. This was clearly seen in their examples in which these pre-service mathematics teachers wrote down or explained the mean calculation to the question on how the average would have been obtained. In fact, two of the pre-service mathematics teachers mentioned in their responses that the mean is the average.

The pre-service mathematics teachers in the present study believed that the average is synonymous to the arithmetic mean similar to the finding of Leavy and

O'Loughlin (2006) which revealed that almost fifty-two percent of the pre-service mathematics teachers' in their study viewed the term average as synonymous to the arithmetic mean. Groth and Bergner (2006) also revealed the same whereby the highest number of respondent explanation in their study related to the mean as the average of a set of data. The responses included the direct mention of the mean as the average and the "add and divide" algorithm of the mean calculation.

On the other hand, the present study found that among all the six pre-service mathematics teachers, only one pre-service mathematics teacher mentioned that the mode can be used as the average. Whereas, another pre-service mathematics teacher mentioned that the mode can be utilised as a guide to decide on the average. The rest totally disagreed on the use of the *mode as the average*. These pre-service mathematics teachers revealed that the average should be based on the mean. They lacked in the content knowledge that average can be referred to mean, median, or mode therefore, limited their substantive knowledge in justifying the use of mode as the average (Jairaman, Sharifah Norul Akmar, & Suzieleez Syrene, 2015). This finding is found to be in contrast to the fact that the mode as a measure of what is typical in a data set and can be a representative of a data set (Groth & Bergner, 2006). This characteristic of the mode can be used as a quick method in reporting an average (Watson, 2006).

In relation to the knowledge of the *median as average*, the finding of this study depicts that none of the pre-service mathematics teachers had mentioned that the median is likely to be the average in the given situation involving human population. Instead the pre-service mathematics teacher who mentioned that the mode can be utilised as the average and the other pre-service mathematics teacher who mentioned that the mode can be utilised as a guide to decide on the average repeated again that the average can be based on the mode for a situation where the median is very likely to be the average.

The remaining four pre-service mathematics teachers mentioned that the average was based on the mean calculation which considered all the data including the outliers. These pre-service mathematics teachers believed that when an average is considered for a population such as Malaysian primary school students, all the data need to be considered in the mean calculation including the outliers.

This finding showed that the pre-service mathematics teachers in the present study had a lack of knowledge on situations that dealt with outliers in human population and the effect of outliers on the averages. They lacked the knowledge that the ideal measure that likely to be the average in situations dealing with human population would be the median. Here, the context along with the discussion on the shape of the distribution and outliers provide an idea on why the median is often the appropriate statistic to use in these kinds of situations dealing with human population (Watson, 2006).

The pre-service mathematics teachers in the present study are not far different from the students in a previous study by Zawojewski and Heckman (1997) which reported that their students were unable to realize that in some contexts the median can be more appropriate measure of centre (Zawojewski & Heckman, 1997).

The pre-service mathematics teachers in the present study were also found to be not much different compared to almost half of 316 prospective teachers in a previous study conducted in Southern Africa who were classified below the basic level of understanding related to the median (Hobden, 2014). These South African teachers' explanations were found to be lacking of the evidence of centre or middle and also the average as an interpretation of the median similar to the pre-service mathematics teachers in the present study.

"Construction problems are very revealing of students' understanding of the relationship between data and average" (Mokros & Russell, 1995, p. 23) and the different distributions that any statistically literate readers of data can see from an average that could be represented by this indicator. The idea of average in this study was explored through these kind of construction problems. Therefore, tasks on data construction based on the given average were used to further investigate the pre-service mathematics teachers' knowledge of the average.

In relation to the *idea of average* both in none restricted and restricted cases, all six pre-service mathematics teachers in this study showed that they held strong to the knowledge of the average as an algorithmic relationship similar to the finding of Russell and Mokros (1991) which mentioned that majority of the teachers in their study also fell in the category of the "average as algorithmic relationship".

The pre-service mathematics teachers in the present study showed manipulation of the mean algorithm similar to the findings of Leavy and O'Loughlin (2006) which reported that thirty-three percent and forty-one percent of their participants utilised the manipulation of the mean algorithm as a strategy to construct data sets for both the none-restricted and the restricted cases respectively. A common response used by the participants in Leavy and O'Loughlin (2006) involved multiplying the average price with the number of elements to find the total was also found common in all the six pre-service mathematics teachers in the present study for both none-restricted and restricted cases.

It was found that five out of the six pre-service mathematics teachers had divided again the total with the 'n' values that summed up to the total for the none-restricted case. They utilised this mean value as the value for all the prices. Thus, this finding showed that these pre-service mathematics teachers also had the knowledge of the average as modal with algorithmic relationship.

However, this finding is found to be slightly in contrast with Mokros and Russell (1995) who studied students between the ages 10 to 12. They found that their students constructed data using an "average" value that was consistently seen as the value that occurs with the greatest frequency and that they rarely attempted to use the mean algorithm. The students in Mokros and Russell (1995) held to the knowledge of average as modal and with no knowledge of algorithmic relationship unlike the pre-service mathematics teachers in the present study who exhibited both knowledge of average as modal and algorithmic relationship.

Meanwhile, four out of the six pre-service mathematics teachers utilised the average as reasonable with algorithmic relationship. One of them utilised this idea for the none-restricted case whereas the remaining three utilised for the restricted case. Although these pre-service mathematics teachers utilised the mean algorithm to guide their construction based on the given average but they scrutinized their choice of values based on reasonableness. They indicated that the average is centred roughly within the rest of the data. This is similar to the finding of Mokros and Russell (1995) which mentioned that some of the children in their study had used average as reasonable as an approach to solve problems.

Only one pre-service mathematics teacher utilised the average as mathematical point of balance combined with algorithmic relationship for both none-restricted and restricted cases. She actually worked with the values in pairs in an earlier construction that she constructed using the average as modal. For each pair of 27, she adjusted the values to other values other than 27 which still averaged out to 27 such as 28 and 26.

She did the same to the other pairs too and tried to avoid 27 in the restricted case. This pre-service mathematics teacher mentioned that she can construct many data sets based on this idea. However, she also utilised the mean algorithm to find the total in which she checked if the total of values that she placed for the prices added to this total.

This pre-service mathematics teacher looked at the whole construction problem into smaller parts. She found the "sub means" as a way to solve the averaging problem. Similar finding was also reported in Leavy and O'Loughlin (2006) where the pre-service teachers in their study had utilised the "set and adjust" method to find the mean in pairs and compared it to the required mean. Values were continuously adjusted until the required mean value was achieved. Mokros and Russell (1995) mentioned that this characteristic is linked to the predominant approach of average as mathematical point of balance.

Thus, it can be concluded that the pre-service mathematics teachers in the present study lacked in the *subject matter knowledge of measures of central tendency as averages*. This might be due to their lack of understanding related to the term "average". They believed that the average is synonymous to the mean similar to the participants in Leavy and O'Loughlin (2006). Thus, when the term average appeared in context, almost all the pre-service mathematics teachers related it to the mean. They were least unaware that the term average can be referred to any of the three measures of centre given to the context of the situation where the term had been used. The pre-service teachers in the present study were unable to discern properly a value which is typical for a given context. Similar to a finding highlighted by a previous study by Batanero, Godino, & Navas (1997).

The pre-service mathematics teachers in the present study also held a very strong procedural knowledge related to the mean calculation and utilised the mean algorithm for almost every average problem that they encountered which supported the claim made by Jacobbe and Carvalho (2011) about teachers exaggerated reliance upon procedural knowledge. Thus, this explains why students' experiences related to this topic are also procedural by nature (e.g. Shaughnessy, 2007, Pollatsek, Lima, & Well, 1981; Mevarech, 1983; Mokros & Russell, 1995; Cai, 1995; Cortina, Saldanha, & Thompson 1999).

One reason behind the pre-service mathematics teachers' lack of the knowledge of averages might be due to their own learning experiences involving the term average which would have been very much linked to the mean and the median or the mode were less focused on. Another reason would have been the lack of exposure on activities involving the average in various contexts and over emphasis on procedures involving the measures of central tendency.

In relation to the pre-service mathematics teachers' subject matter knowledge of measures of central tendency as data representation, it can be concluded that all six pre-service mathematics teachers had very limited knowledge of the *mode as data representation*. None of them have actually utilised the mode as a form of data representation with proper justification towards its use (Jairaman, Sharifah Norul Akmar, & Suzieleez Syrene, 2016a, 2016b).

Although, four of the pre-service mathematics teachers mentioned that the mode can be used as a form of data representation but they have justified its use inappropriately. One of the pre-service mathematics teacher justified that the mode can be a form of data representation because it represented the "majority". However, she also explained that the mode can be a form of data representation only if the mode is close to all other data points in the data set. If there exist one or two extreme values, she claimed that the mode cannot be utilised as a form of data representation.

Whereas, the other three pre-service mathematics teachers justified that the mode can be utilised as a form of data representation only if the mode value is around the average value or if the mode can be counted. They held to the knowledge that the mode is numerical all the time similar to the findings of Groth and Bergner (2006) whereby most of the pre-service elementary teachers in their study also implied that the mode is only of use for numerical data sets. Here, the pre-service mathematics teachers in the present study appeared to have a lack in the knowledge of the mode which can also be categorical.

Meanwhile, two of the pre-service mathematics teachers totally disagreed that the mode can be utilised as a form of data representation. They said that the mode is based on one particular data which carried the highest frequency, therefore cannot be utilised as a form of data representation (Jairaman, Sharifah Norul Akmar, & Suzieleez Syrene, 2016a, 2016b). This finding contradicted with the fact that the mode is one of the measures of centre which attempts to find a "typical" amount that summarise a certain data set (Jacobbe & Carvalho, 2011). Therefore, the mode can be a representative of a data set from which it arises (Groth & Bergner, 2006).

In relation to the decision on the *best data representation measure*, it is important to understand the distinguishing features of the arithmetic mean and the median in order to appreciate why it is important to use one or the other (Watson, 2006). The present study found that three out of the six pre-service mathematics teachers chose the median as the best data representation measure based on the influence of the extreme and outlying values.

Among the three pre-service mathematics teachers, two of them also mentioned that besides the median, the mean which is calculated after excluding the outlier can also be considered as the best data representation measure. Whereas another two pre-service mathematics teachers from the remaining three pre-service mathematics teachers mentioned that the mean which is calculated after excluding the extreme value is the best data representation measure. These teachers did not mention anything on the median being a more robust measure to represent the data in the presences of an outlier.

Most pre-service mathematics teachers in the present study were aware about the nature of bias introduced by an outlying value on the mean. They either chose the median or the mean which was calculated after excluding the extreme data as the best data representation measure for the data sets that contained extreme values. This finding was found to be similar to a previous finding by Strauss and Bichler (1988) where the students in their study were aware that particular data values can influence the mean.

However, there was one pre-service mathematics teacher in the present study who chose the mean which was calculated based on all the data including the extreme value as the best data representation measure. This pre-service mathematics teacher lacked the knowledge on the effect of outliers on the mean similar to the finding of Batanero et al (1997) which reported that thirty-four percent out of 273 pre-service teachers' in their study showed little or no understanding on the effect of outliers on the mean. Estrada et al (2004) also reported that many pre-service primary teachers in their study were not conscious on the effect of outliers on the mean, and Pollatsek, Lima and Well (1981) also found that the college students in their study did not take the extreme value into account for a sample mean. The ignorance on the effect of the outliers on the mean can create tremendous bias issue (Watson, 2006).

The discussion related to the robustness of the measures of central tendency and deciding on which of the three measures of central tendency is least to be biased by extreme or outlying data points along with the likelihood of this measure in representing the data is important in statistical literacy (Watson, 2006). It was found that four out of the six pre-service mathematics teachers were aware that the median is a more adequate measure than the mean to represent a set of data in the presences of an extreme data point.

They also claimed that the mean can also be considered if the mean is calculated after excluding any extreme value. In fact, one of the pre-service mathematics teacher mentioned that the mean which is calculated after excluding the extreme value is more adequate to represent a set of data with an extreme point with no mention at all about the median.

The pre-service mathematics teachers in the present study were aware that an extreme data can influence the mean similar to the finding of Strauss and Bichler (1988) which reported that the students in their study were quite aware of the influence of an extreme data on the mean. However, this finding contradicted the claim of one pre-service mathematics teacher whom calculated the mean based on all the data including the extreme as the most adequate measure to represent a set of data that contained extreme point similar to Pollatsek, Lima, and Well's (1981) finding where by the college students in their study did not take into account the extreme value for a sample mean.

Therefore, in relation to the *subject matter knowledge of measures of central tendency as data representation*, it can be concluded that most pre-service mathematics teachers in the present study had sufficient knowledge related to the *robustness of the measures of central tendency* and the *best data representation measure*. Most of the pre-service mathematics teachers were aware of the influence of an outlier on the mean and some of them were also aware that the median can be a better data representation measure for a set of data that contained an outlier. However, the pre-service mathematics teachers in the present study seem to have a lack in the knowledge related to the *mode as data representation*. They were not aware that the mode can be utilised as a form of data representation which might be due to the lack of attention given to this concept in most mathematics curriculum especially involving this concept in various representations and contexts.

With regard to the *subject matter knowledge of measures of central tendency procedures*. The findings of the present study found that all six pre-service mathematics teachers had sufficient knowledge related to the *idea of mode* for numerical data without the task explicitly stating to do so. They utilised the frequency tally to obtain the frequencies of each category and identified the mode based on the category with the highest frequency. The pre-service mathematics teachers in the present study had no issue in identifying the mode based on the corresponding category which had the highest frequency which contradicted with the findings of Fernandes (2009, as cited in Santos & da Ponte, 2013) where some of his participants selected the biggest frequency instead of the corresponding value of the variable.

However, when these pre-service mathematics teachers were presented with another task that carried data in tabular form with one none-numerical category in the table. Only one out of the six pre-service mathematics teachers could identify the mode correctly based on the "most frequent number" or "the number with the highest frequency" within the data set. Three of them showed some knowledge that the mode can be categorical which is in contrast to Fernandes (2009, as cited in Santos & da Ponte, 2013) whereby some prospective teachers in his study failed to select the corresponding value of the variable for the selected biggest frequency. However, the pre-service mathematics teachers in the present study had a misconception that the biggest number of comments made on that particular day would indicate the most frequent number of comments.

Whereas, two of the pre-service mathematics teachers had limited knowledge on how to identify the mode because they had actually taken the biggest number of comments (a misconception that this value is indicating the most frequent number of comments) as the mode. This finding is found to be similar to Martins, Pires and Barros (2009, as cited in Santos & da Ponte, 2013) where their participants have interpreted the mode as "the biggest number on the results table" which was also similar to Santos and da Ponte (2013) which reported one of the problematic interpretations of the mode by their participants was the mode as the "highest value".

For these pre-service mathematics teachers, one of the defining characteristics of the mode is a "number" similar to Groth and Bergner (2006) in which their study reported that the pre-service elementary teachers mostly implied that the mode is only of use for numerical data sets when the mode is not necessarily numerical all the time and can be categorical in certain situation. Nevertheless, Groth and Bergner's (2006) study was unable to distinguish clearly the characteristic of the mode which could be a category as well as a number in their participants.

However, this study had investigated on the idea of mode which can be a category as well as a number. The finding depicts that some of the pre-service mathematics teachers in the present study knew that the mode can be a category as well as a number (Jairaman, Sharifah Norul Akmar, & Suzieleez Syrene, 2016a, 2016b) which fills the gap in a previous study by Groth and Bergner (2006) but the participants in the present study showed a lack of knowledge of the mode as "the most frequent number" or "the most common number" for data presented in tabular form (Jairaman, Sharifah Norul Akmar, & Suzieleez Syrene, 2016a, 2016b).

The findings of the present study also had identified several interpretations of the concept of the mode similar to Martins, Pires and Barros (2009, as cited in Santos & da Ponte, 2013) such as “the mode to the biggest number on the table”, “the value that appears more times” and “the biggest frequency category or interval”.

In relation to the *idea of median*, five out of six pre-service mathematics teachers in the present study carried the correct median procedure both for even and odd numbered data. However, only one out of the five pre-service mathematics teachers mentioned that the data can be ordered in both ascending and descending order before the median is obtained. The remaining four mentioned that the data should be arranged in ascending order only.

This finding is found to be in contrast to the finding of Jacobbe (2012) which reported that the elementary school teachers in his study showed a lack of procedural knowledge of finding a median. The teachers in Jacobbe (2012) had confused that the algorithm for calculating the mean with the median procedure and found the median from an unordered data similar to Martins et al. (2009, as cited in Santos & da Ponte, 2013) which also reported that the most common mistake in finding the median among the participants in their study was to determine the central value without ordering the data first. However, this was not the case for most of the pre-service mathematics teachers in the present study.

Nevertheless, one of the pre-service mathematics teacher in this study showed a rather different procedure of finding the median. She ordered the data in ascending order first than she added all the data and divided the total by two. The value that she obtained she used to locate the position of the median by adding the data in the rearrangement one by one until she found the location of the value in which she believed to be the location of the median. If this value fell in between two different numbers, she added the two numbers and then divided by two to find the median.

This error suggests that this pre-service mathematics teacher showed a lack in the procedural knowledge in finding the median in which might have been caused by a short-circuit in her thinking due to the introduction of an algorithm similar to what was claimed by Russell and Mokros (1991). Russell and Mokros (1991) mentioned that the introduction of an algorithm as a procedure disconnected from students' informal understanding of middle which causes a short-circuit in the reasoning.

The mean is not just a numeric algorithm-based summary of a set of numbers but it also has representational attributes. The representative nature of the mean can be considered in making sense of why one would want to combine data sets to obtain a weighted mean. Even though knowing the algorithms are important but to understand the connections between the algorithms to the purposes of using summary statistics are equally important too (Watson, 2006). In regards to the pre-service mathematics teachers' knowledge of solving weighted mean problems, it was found that all six pre-service mathematics teachers had sufficient knowledge to do so.

These pre-service mathematics teachers had utilised the *backward mean calculation* to obtain the totals for the groups from the group mean and they had no problem in doing so which contradicted the finding of Cai (1995) that reported although most students could calculate a mean but they had great difficulty in working backwards. However, the finding of this study is supported by the finding of another recent study done on 166 first year pre-service teachers in Greece which reported that eighty-three percent of their participants were able to reverse the algorithm for the computation of the mean (Koleza & Kontogianni, 2013).

All six pre-service mathematics teachers in the present study were aware of the *representative nature of the mean* in making sense the need for data sets to be combined in a weighted mean problem. After when they combined these group totals and carried out the *forward mean calculation* to obtain a weighted mean. This finding is found to be contradicting the findings of previous studies such as Pollatsek, Lima and Well (1981) which found that most undergraduate college students in their study were unable to correctly weight and combine two means into a single mean. Leavy and O'Loughlin (2006) also found that only twenty-one percent of the pre-service teachers in their study correctly solved the weighted mean problems.

The pre-service mathematics teachers in the present study were aware of what the overall mean is supposed to represent rather than thinking of the mean as solely the result of the application of a computational rule which contradicted the finding by Pollatsek, Lima and Well (1981) where most of their participants saw the mean as a computational rule rather than a conceptual act. These previous studies (Pollatsek, Lima & Well, 1981; Leavy & O'Loughlin, 2006) also reported that their participants tended to weight the samples equally or found the mean of the given means.

Thus, in relation to the *subject matter knowledge of measures of central tendency procedures*, it can be concluded that almost all the pre-service mathematics teachers in this study had sufficient knowledge except for identifying the mode based on data given in tabular form. All the pre-service mathematics teachers had no problem in carrying out the procedures in calculating the mean. Almost all of them could carry out the procedures to find the median correctly too and all of them had no issues in identifying the mode for numerical data.

However, in relation to identifying the mode for data in tabular form, almost all the pre-service mathematics teachers had problem. They had issues related to identifying the mode based on "the most frequent" or "the highest frequency" and also had issues related to identifying the mode based on the category with "the highest frequency" or "the most frequent" category which might be due to the lack of attention given to the mode in most mathematics curriculum especially involving different forms of data and the presentation of this concept in different contexts.

The pre-service mathematics teachers in the present study held strong knowledge of procedures similar to the rest of teachers in previous studies (Jacobbe & Carvalho, 2011). This explains that these teachers' experiences involving the measures of central tendency might have been procedural and that is why these teachers held exaggerated reliance upon procedural knowledge in which indirectly influenced their teachings of these measures leading to students' experiences to be procedural too (e.g. Shaughnessy, 2007, Pollatsek, Lima, & Well, 1981; Mevarech, 1983; Mokros & Russell, 1995; Cai, 1995; Cortina, Saldanha, & Thompson 1999).

With regards to the *subject matter knowledge of measures of central tendency as data summary*, all six pre-service mathematics teachers in this study had *utilised all or any one of the measures of central tendency to summarise both equal and unequal sized numerical data* without specifically told. Five out of six pre-service mathematics teachers in this study had also *summarised equal and unequal sized graphical data using measures of central tendency*, without specifically told.

However, the remaining one pre-service mathematics teacher did not use the measures of central tendency to summarise unequal sized graphical data. He explained that since the graphical data was unequal sized, therefore he claimed that the measures of central tendency cannot be utilised to make any judgement about the data and there is no need to summarise the data using these measures.

Thus, in relation to the *subject matter knowledge of measures of central tendency as data summary* it can be concluded that almost all pre-service mathematics teachers utilised the measures of central tendency as a data summary tool without the tasks explicitly stating to do so. The pre-service mathematics teachers were aware that measures of central tendency are effective data summary tools. The present study found that research related to utilising the measures of central tendency as a data summary tool without the tasks explicitly stating to do so has not been explored much in the past research when this is an important aspect in the enhancement of statistical literacy.

One of the goals of statistics related to the enhancement of statistical literacy is to utilise the measures of central tendency as a significant tool to compare and make judgements about data sets without being specifically told (Watson, 2006). In relation to the subject matter knowledge of measures of central tendency as data comparison, five out of the six pre-service mathematics teachers had utilised the mean to *compare and draw conclusions about equal and unequal sized numerical data sets*. The remaining one pre-service mathematics teachers had utilised the mode to *compare and draw conclusions about equal and unequal sized numerical data.*

The pre-service mathematics teachers who utilised the mean to *compare and draw conclusions about equal and unequal sized numerical data* have demonstrated the functional knowledge of the mean where the mean is a representative value that can be used to index and characterize a body of data. They utilised the mean as the most efficient measure in comparing two data sets. This finding is similar to Leavy and O'Loughlin (2006) that reported that three quarter of the pre-service teachers in their study also had utilised the mean to compare and draw judgements about unequal sized data sets.

However, the finding contradicted with the finding of a previous study done on middle grade students which found that these students usually did not use the mean when asked to compare data sets (Gal, Rothschild, & Wagner, 1989, 1990). Konold, Pollatsek, Well, and Gagnon (1997) reported similar results with their upper secondary students who also neglected the mean as a comparative measure.

However, when the pre-service mathematics teachers were asked to *compare and draw conclusions on the equal sized graphical data*, half of the pre-service mathematics teachers chose to compare and draw conclusions based on the mode. Whereas, the other half compared and drew conclusions based on the mean. This finding is in contrast with Watson and Moritz (2000) which found that in the assigned tasks of comparing two data sets in graphical form, students often compared the data sets by totaling or by making visual comparisons, rather than by employing the arithmetic mean or any other measure of average.

On the other hand, when these pre-service mathematics teachers *compared unequal sized graphical data*, four out of the six pre-service mathematics teachers preferred to use the mean to compare and draw conclusions based on unequal sized graphical data, one pre-service mathematics teacher chose to compare and draw conclusion based on the mode, and the remaining one did not utilise any of the measures of central tendency. Most of the pre-service mathematics teachers in the present study utilised the mean to compare and draw conclusions based on unequal sized graphical data which is found to be slightly in contrast with the finding of a previous study done on students which found that the mean was least preferred when comparing graphs of two data sets of different sizes (Watson & Moritz, 2000).

Whereas, the remaining one pre-service mathematics teacher whom did not use any of the measures of central tendency had actually compared and drew conclusions about the unequal sized graphical data using visual comparison and also compared the skewness of the graphs. Watson and Moritz (1999) reported that in their investigation on Grade 3 to 9 students' strategies when comparing pairs of data sets both equal and unequal sized in graphical form, visual strategies were one of the strategies used by their students besides comparison of individual data values, to calculating total values, and calculating the means of the groups.

Thus, in relation to the *subject matter knowledge of measures of central tendency as data comparison* it can be concluded that almost all pre-service mathematics teachers utilised the mean as the most efficient measure in comparing two equal and unequal sized numerical or graphical data without the tasks explicitly stating to do so. Some of the pre-service mathematics teachers had utilised the mode to compare and draw conclusions based on equal and unequal sized numerical or graphical data although this measure of central tendency might not be the most efficient measure in comparing two data sets.

Only one pre-service mathematics teachers utilised the visual comparison to compare two unequal sized graphical data instead of the measures of central tendency because he justified that for unequal sized graphical data, the measures of central tendency are not appropriate to compare and draw conclusions.

It can be concluded that some of the pre-service mathematics teachers in the present study lacked in the knowledge of the most efficient measures of central tendency in comparing two data sets, although majority have utilised the measures of central tendency as a tool to compare and draw conclusions about data sets. Most pre-service mathematics teachers used the mean as the functional tool to draw conclusions involving two equal and unequal sized numerical or graphical data. The mode or the median were least preferred. These pre-service teachers were aware that mean takes count of a situation on deciding on the class or the player that or who performed better.

The finding of this study showed that with regard to the overall level of the subject matter knowledge of measures of central tendency based on the emerged themes, only one out of the six pre-service mathematics teachers secured an overall high level whereas the remaining five secured an overall medium level.

The finding depicts that in relation to the subject matter knowledge of measures of central tendency as averages, two out of the six pre-service mathematics teachers secured a high level whereas the remaining four secured a medium level. It was found that the pre-service mathematics teachers who failed to attain a high level was due to their lack in the knowledge of the mode as average and the median as average. These pre-service mathematics teachers lacked the knowledge that the mode is a measure of what is typical in a data set and can be representative of a data set (Groth & Bergner, 2006) whereby this characteristic of the mode can be utilised as a quick method in reporting an average (Watson, 2006).

The pre-service mathematics teachers also lacked in the knowledge that the median is likely to be the average in data involving human population. Moreover, most of the pre-service mathematics teachers believed that the average is synonymous to the mean regardless the context of the given situation. Thus, they conclude that an average given in any situation has to be based on the mean. In addition, the pre-service mathematics teachers also exhibited strong procedural knowledge of the mean and utilised the mean algorithm for the data construction tasks for a given average value which clearly indicated that their knowledge of the average is closely tied to the mean.

These findings on the level of subject matter knowledge of measures of central tendency as averages showed that the pre-service mathematics teachers in the present study are still unable to appreciate the nuances of the usage of central tendency terms in authentic contexts which is regarded as an important aspect related to statistical literacy. Watson (2006) highlighted this at the second tier of her Statistical Literacy Hierarchy.

With regard to the pre-service mathematics teachers' level of subject matter knowledge of measures of central tendency as data representation, two of the pre-service mathematics teachers attained a high level, three attained a medium level and one attained a low level. It was found that the pre-service mathematics teachers who failed to attain a high level was due to their lack in the knowledge of the mode as data representation. These pre-service mathematics teachers either mentioned that the mode cannot be utilised as a form of data representation or if they mentioned that the mode can be a form of data representation, they were unable to justify their answers.

With regard to the pre-service mathematics teachers' level of subject matter knowledge of measures of central tendency procedures, five out of six pre-service mathematics teachers attained a high level while the remaining one achieved a medium level. Although most of the pre-service mathematics teachers achieved a high level for the subject matter knowledge of measures of central tendency procedures but it was found that almost all except for one lacked in the knowledge of identifying the mode for data presented in tabular form.

Most of them either identified the mode as the "biggest number" or "the category that referred to the biggest number" which might be due to a misconception that the "biggest number" was actually the "highest frequency". Furthermore, the only pre-service mathematics teacher who attained a medium level for this theme also lacked the knowledge related to the idea of median whereby she had carried out the median procedure incorrectly.

For the subject matter knowledge of measures of central tendency as data summary, five out of six pre-service mathematics teachers attained a high level and the remaining one attained a medium level. Again, the pre-service mathematics teacher who attained a medium level for this theme was because of her knowledge of summarising the given data using median was incorrect because she carried out the median procedure incorrectly.

Finally, for the subject matter knowledge of measures of central tendency as data comparison, half of the pre-service mathematics teachers had utilised the appropriate measure to compare and draw conclusions about two equal or unequal sized numerical or graphical data. Therefore, three pre-service mathematics teachers attained a high level for this theme. On the other hand, two pre-service mathematics teachers who attained a medium level was because these teachers lacked in the knowledge of utilising the appropriate measure to compare and draw conclusions about two equal or unequal sized graphical data.

The one pre-service mathematics teacher who attained a low level for this theme lacked in the knowledge of utilising the appropriate measure to compare and draw conclusions about equal or unequal sized numerical or graphical data. This pre-service mathematics teacher had utilised the mode instead of the mean.

Lastly, this section is concluded with a comparison on how the pre-service mathematics teachers in the present study differ from the pre-service mathematics teachers in previous studies. The pre-service mathematics teachers in the present study showed not much issues in relation to their procedural knowledge of measures of central tendency. All of them carried out the algorithm for calculating the arithmetic mean properly which was not the case with the pre-service mathematics teachers in Batanero, Godino, and Navas (1997).

Almost all of them carried out the median procedures correctly and showed reasonable understanding of the median unlike the prospective teachers in Hobden (2014). The pre-service mathematics teachers in this study also had no issues in inverting the mean algorithm and carried out the weighted mean problems correctly which was not the case with the teachers in Estrada, Batanero, and Fortuny (2004) and Callingham (1997).

These findings strongly indicate that the learning experiences of the pre-service mathematics teachers in the present study might have emphasized on the procedures of measures of central tendency which explains why these teachers held a better level of subject matter knowledge of measures of central tendency for the themes namely; procedures and data summary.

The pre-service mathematics teachers in the present study also showed reasonably good understanding of the effect of the outliers on mean unlike the pre-service teachers in previous studies by Batanero, Godino, and Navas (1997) and Estrada, Batanero, and Fortuny (2004). However, the pre-service mathematics in this study had issues in discerning properly that a value which is a typical for a given context similar with the pre-service teachers in Batanero, Godino, and Navas (1997) and they carried out routine application of the mean algorithm without taking into account the context similar to Estrada, Batanero, and Fortuny (2004)

These findings explain why their levels of the subject matter knowledge of measures of central tendency for the themes data representation and data comparison were lower than the other themes. The pre-service mathematics teachers in this study lacked in the knowledge of measures of central tendency with reference to context which called for attention in the teaching and learning of statistics.

Furthermore, the pre-service mathematics teachers in this study showed little variation in relation to solving problems related to the mean which was in contrast to the pre-service mathematics teachers in Gfeller, Niess, and Lederman (1999) who held multiple views of the mean and a deeper conceptual understanding of the mean. In relation to the average, Russell and Mokros (1991) mentioned that majority of the teachers in their study fell in the category of "average as an algorithm relationship" and showed varying levels of success with their application of the algorithm in finding the mean value which was also the case with the pre-service mathematics teachers' in the present study.

However, most of the pre-service mathematics teachers in the present study also exhibited average as reasonable sense which was not exhibited by the teachers in Russell and Mokros (1991). Yet, the pre-service mathematics teachers in present study viewed "mean" as synonymous to "average" similar with the participants in Leavy and O'Loughlin (2006).

Thus, explaining why the pre-service mathematics teachers' levels of subject matter knowledge of measures of central tendency as averages were lower compared to certain themes. These findings indicate that emphasis is needed on the term average in relation to the three measures of central tendency for the enhancement of statistical literacy in the teaching and learning of statistics.

Thus, in conclusion the present study has extended on the study related to the subject matter knowledge of measures of central tendency, namely for the enhancement of statistical literacy in which Gal (2002) have highlighted as lacked in attention. The findings of the present study also support the general assumption that the pre-service mathematics teachers who come into the teaching profession might not be well equipped with the subject matter knowledge (Even, 1993; Nik Azis, 2003).

Five main themes that were identified as the subject matter knowledge of measures of central tendency involved in the enhancement of statistical literacy namely; averages, data representation, procedures, data summary, and data comparison extend to the content needed in the preparation of teachers and students of the concepts of central tendency for everyday life situations.

Implications of Findings

Findings of this study have several implications for the pre-service mathematics teacher education, curriculum development, text book and statistical literacy.

Implications to pre-service mathematics teacher education. In general, the findings of this study suggest that one can identify two groups of pre-service mathematics teachers.

The first group composed of pre-service mathematics teachers who was seen to have an overall high level of subject matter knowledge of measures of central tendency. Whereas, the second group comprised of pre-service mathematics teachers who appeared to have an overall medium level of subject matter knowledge of measures of central tendency. However, based on the sample of this study, the third group which supposedly consist the pre-service mathematics teachers with low overall level of subject matter knowledge of measures of central tendency was not seen.

The implication of this finding is that the mathematics teacher educators now have the challenge to develop a mathematics teacher program that meets the need of all these groups. This mathematics education program is expected to help our pre-service mathematics teachers to construct their subject matter knowledge of measures of central tendency in terms of the five identified themes, namely averages, data representation, procedures, data summary, and data comparison. Eventually this will enhance the pre-service mathematics teachers' level of subject matter knowledge of measures of central tendency.

To be specific, such mathematics teacher program is aimed to help our pre-service mathematics teachers to understand the importance of these five themes for the teaching and learning of the enhancement of statistical literacy. This would also facilitate our pre-service mathematics teachers to include the necessary components involved in the five themes appropriately in their lesson planning and teaching.

In this study, clinical interview technique was used to elicit pre-service mathematics teachers' subject matter knowledge of measures of central tendency and such technique had enabled the researcher to elicit pre-service mathematics teachers' subject matter knowledge of measures of central tendency in terms of the four constructs, namely with reference to the context, in handling bias, in problem solving, and in making inference. Such technique is also recommended to be used to elicit pre-service mathematics teachers' subject matter knowledge of other statistics topics involved in the enhancement of statistical literacy.

In relation to the data collection, it is clear that the method can be applied by mathematics teacher educators to proffer enriching experiences for their pre-service mathematics teachers. Similarly, the clinical interview tasks devised for the purpose of this study can also be utilised by mathematics teacher educators to investigate the nature and level of their pre-service mathematics teachers' subject matter knowledge of measures of central tendency.

From the findings of this study, it was apparent that all the six pre-service mathematics teachers in this study attained a medium level of subject matter knowledge of measures of central tendency with reference to context. The implication of this finding is that mathematics teacher educators need to organize teaching and learning activities that provide experiences for our pre-service mathematics teachers with the concept of measures of central tendency in the context of everyday life.

Implications to curriculum development. Almost all the pre-service mathematics teachers in this study had the knowledge that the average is synonymous to the mean. They thought that when average is mentioned than it should be referring to the mean.

However, in relation to statistical literacy the general term average can mask any of the three measures of central tendency. The implication of this finding is that the curriculum specifications should be clear with the terms of measures of central tendency given in the context of daily life situations.

Thus, the mathematics teachers need to organize teaching and learning activities following these curriculum specifications that provide experiences for the students with the concept of the average in the context of everyday life. Through such exposures, the knowledge that the word average is a reflection of central tendency and that all three measures of central tendency; the mean, the median, and the mode contribute to the idea of average can be developed.

The finding of this study showed that the pre-service mathematics teachers lacked in the knowledge to identify the mode correctly based on the most frequent or highest frequency for certain situation such as when dealing with data presented in tabular form. Some lacked at the knowledge that the mode is not necessarily numerical all the time and can be categorical and some had a misconception that the “biggest number” would indicate the “highest frequency”.

Apparently, this lack also indicated a lack in their knowledge of the mode as a representative of a data set because the mode can describe a typical behaviour of a data set and therefore it can be used as a form of data representation. This characteristic of the mode is also used as a quick method in reporting central tendency or an average.

The implication of this finding is that the curriculum specification should clearly state to introduce the concept of the mode in daily life context using data presented in various forms such as tabular, numerical, and graphical.

The mathematics teachers need to incorporate teaching and learning activities in the classroom that provide opportunities for the students to identify the mode based on these various contexts. They must also include the discussion on the characteristic of the mode in relation to the measures of central tendency including the connection of this concept to the term average in everyday life.

The findings of this study showed that some of the pre-service mathematics teachers were unable to identify the best representation measure based on the given data or were unable to justify appropriately the reason for choosing such measure as the best measure to represent a set of data. The implication of this finding is that the curriculum specification should clearly state to incorporate the ideas of the measures of central tendency in terms of handling bias.

The mathematics teachers need to incorporate teaching and learning activities in the classroom that provide opportunities to include the discussions of the robustness of the measures of central tendency in terms of handling bias.

The finding of this study showed that some of the pre-service mathematics teachers lacked in the knowledge of utilising the appropriate measure of central tendency to compare and draw conclusions based on equal and unequal sized data. Some also had difficulty to obtain the measures of central tendency when the data was presented in graphical form.

The implication of this finding is that the curriculum should incorporate the ideas of measures of central tendency in relation to data summary.

The mathematics teachers need to incorporate teaching and learning activities in the classroom that provide opportunities for students to utilise measures of central tendency to summarise data presented in groups of various representations such as numerical, tabular or graphical and later compare and draw conclusions based on these groups utilising the most appropriate measure of central tendency.

All the above findings and implications are all in line with the recommendation of the Malaysian Mathematics Curriculum Specifications which suggests that the mathematics educators need to provide opportunity for students to apply appropriately the concepts of central tendency to interpret and solve problems in their daily life situations.

Implications to text book. The findings of this study showed that the subject matter knowledge of measures of central tendency involved in the enhancement of statistical literacy requires beyond procedural knowledge.

It requires more than knowing how to compute the mean, how to identify the mode and how to locate the median. Thus, the implication of this finding is that the school text books should also incorporate activities that goes beyond calculation and procedures to find the measures of central tendency. Activities that include measures of central tendency in the five emerged themes namely; averages, data representation, procedures, data summary, and data comparison using real life situations should be incorporated.

Implications to statistical literacy. The findings of this study showed the extension in the knowledge of measures of central tendency. However, the present study focused on the aspects of measures of central tendency involved in the enhancement of statistical literacy which was found to be lacking. The implication of this finding is that similar research is recommended to be carried out to elicit pre-service mathematics teachers' subject matter knowledge of other statistical topics involved in the enhancement of statistical literacy. Thus, to advance the knowledge component involved in the enhancement of statistical literacy.

Recommendations for Further Research

The followings are the recommendations for the extension of this study and further research. This study only involved six pre-service mathematics teachers. The subjects were drawn from the pre-service mathematics teachers who enrolled in the 4-year Bachelor of Science Education (B. Sc. Ed.) program in a public university. Therefore, it is recommended that the present study to be extended to other pre-service mathematics teachers who enrolled in the 4-year Bachelor of Science Education (B.Sc. Ed.) program in this public university, in similar program in other public universities, and teachers training institutes so as to verify and elaborate the findings of the present study.

This study focused on pre-service mathematics teachers' subject matter knowledge of measures of central tendency. Therefore, it is recommended that the study to be extended to in-service mathematics teachers and school students at various levels. This may contribute towards a wider knowledge base of the teachers' and students' subject matter knowledge of measures of central tendency involved in the enhancement of statistical literacy.

This study focused on pre-service mathematics teachers' subject matter knowledge of measures of central tendency involved in the enhancement of statistical literacy. Thus, it is recommended that further research examine pre-service mathematics teachers' subject matter knowledge of other statistical topics involved in the enhancement of statistical literacy. The findings of such studies may contribute towards a wider knowledge base of the teachers' subject matter knowledge of statistical topics involved in the enhancement of statistical literacy.

This study examined the nature and levels of pre-service mathematics teachers' subject matter knowledge of measures of central tendency. The present study did not examine pre-service mathematics teachers' beliefs about measures of central tendency. Thus, it is recommended that further research examine pre-service mathematics teachers' beliefs about measures of central tendency as well as their beliefs about teaching and learning of measures of central tendency.

Contributions of the Study

Studies related to measures of central tendency mostly have focused on students and only limited studies have focused on teachers (Jacobbe & Carvalho, 2011). Moreover, there existed only a handful of studies which focused on teachers' professional knowledge (e.g. Cai & Gorowara, 2002; Olfos & Estrella, 2010; Watson & Callingham, 2013). Thus, the present study contributes towards the extension on studies related to teachers.

Most of the studies on measures of central tendency have focused on aspects such as; understanding of the average and its properties (Mokros & Russell, 1995; Strauss & Bichler, 1988), weighted mean (Pollatsek, Lima, & Well, 1981; Mevarech, 1983), and arithmetic mean (Batanero, Godino, & Navas, 1997; Callingham, 1997; Gfeller, Niess, & Lederman, 1999; Leavy & O'Loughlin, 2006).

Among the three measures of central tendency, the general concept of average and the mean had been widely focused in previous studies. The median had a reasonable amount of focus but the mode had the least attention. Limited studies have actually focused on all three measures of central tendency (e.g. Groth & Bergner, 2006; Jacobbe, 2007, 2008).

Therefore, the present study contributes towards the extension on studies related to all three measures of central tendency and has given focus on the mean, the median, the mode, and the average. The present study contributes to the knowledge extension on the concept of the mode which was found to be scarce. Moreover, the present study also shed some light on the term average which can mask any of the three measures of central tendency in relation to statistical literacy which was focused in previous studies on the general concept of average.

The present study also contributes to the extension of the subject matter knowledge of measures of central tendency involved in the enhancement of statistical literacy needed by prospective teachers for the preparation of these measures to everyday life in which the ideas of these measures related to daily life have not received much attention.

The findings of the present study through the emerged themes have identified some of the subject matter knowledge of measures of central tendency that are important for the enhancement of statistical literacy. The present study went beyond the aspect of procedures involved in the concepts of measures of central tendency and investigated on the usage of these concepts in everyday life. Thus, through these findings some of the ideas of measures of central tendency required for everyday life were identified.

The present study also contributes towards knowing if the pre-service mathematics teachers through the selected sample are well prepared in terms of the subject matter knowledge of measures of central tendency. The study helped to understand if these pre-service mathematics teachers have adequate subject matter knowledge of measures of central tendency and if they are well equipped in terms of these concepts.

Gal (2002) highlighted on the lack of attention given on the nature of the knowledge involved in the development of statistical literacy in various life context. Therefore, the present study contributes to the gap in the nature of the knowledge involved in the development of statistical literacy.

The present study's framework was based on Gal's (2002) model of statistical literacy where the investigation of the subject matter knowledge of measures of central tendency was focused on. The findings of the present study revealed the subject matter knowledge of measures of central tendency needed for the enhancement of statistical literacy.

The five themes identified through the findings of the present study; namely averages, data representation, procedures, data summary, and data comparison contributes towards the basic ideas for further research in the topic of measures of central tendency.

References

Ball, D. L. (1988). *The Subject Matter Preparation of Prospective Mathematics Teachers: Challenging the Myths.* National Center for Research on Teacher Education, East Lansing, MI.

Ball, D. L. (1990). The mathematical understandings that prospective teachers bring to teacher education. *The elementary school journal*, 449-466.

Ball, D.L., & Bass, H. (2000). Interweaving content and pedagogy in teaching and learning to teach: Knowing and using mathematics. In J. Boaler (Ed.), *Multiple perspectives on teaching and l earning of mathematics* (pp. 83-104).

Ball, D. L., & McDiarmid, G. W. (1989). The subject matter preparation of teachers. In W.R. Houston (Ed.). *Handbook for Research on Teacher Education.* New York: Macmillan.

Ball, D. L., Hill, H. C., & Bass, H. (2005, fall). *Knowing mathematics for teaching: Who knows mathematics well enough to teach third grade, and how can we decide?* American Educator, 14-46.

Ball, D. L., Lubienski, S. T., & Mewborn, D. S. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. In V. Richardson (Ed.). *Handbook of research on teaching* (4^{th} ed., pp. 433-456). Washington, DC: American Education Research Association.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59* (5), 389-407.

Barrett, J. E., Clement, D. H., Klanderman, D., Pennisi, S.J., Polaki, M. V. (2006). Students' coordination of geometric reasoning and measuring strategies on a fixed perimeter task: Developing mathematical understanding of linear measurement. *Journal for Research in Mathematics Education, 37(2)*, 187-222.

Batanero, C., Cobo, B., & Díaz, C. (2003). Assessing secondary school students' understanding of averages. In *Proceedings of CERME* (Vol. 3). Retrieved from http://www.dm.unipi.it/~didattica/CERME3/proceedings/Groups/TG5/TG5_batanero_cerme3.pdf

Batanero, C., Godino, J., & Navas, F. (1997). Concepciones de maestros de primaria en formción sobre promedios (Primary school teachers' conceptions on averages). In H. Salmerón (Ed.), *Actas de las VII Jornadas LOGSE: Evaluación Educativa* (pp 310-340). Granada, Spain: University of Granada.

Baturo, A., & Nason, R. (1996). Student teachers' subject matter knowledge within the domain of area measurement. *Educational Studies in Mathematics*, *31*(3), 235-268.

Begg, A., & Edwards, R. (1999). Teachers' ideas about teaching statistics. *Proceedings of the 1999 Combined Conference of the Australian Association for Research in Education and the New Zealand Association for Research in Education,* Melbourne: Australian Association for Research in Education.

Ben-Zvi, D., & Garfield, J. (2008). Introducing the emerging discipline of statistics education. *School Science and Mathematics*, *108*(8), 355-361.

Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy.* Academic Press: New York.

Biggs, J. B., & Collis, K. F. (1991). Multimodal learning and the quality of intelligent behaviour. In H. A. H. Rowe (Ed.), *Intelligence: Reconceptualization and measurement* (pp. 57-76). Hillsdale, NJ: Lawrence Erlbaum.

Callingham, R. (1997). Teachers' multimodal functioning in relation to the concept of average. *Mathematics Education Research Journal*, *9*(2), 205-224.

Cai, J. (1995). Beyond the computational algorithm: Students' understanding of the arithmetic average concept. In L. Meira & D. Carraher (Eds.), *Proceedings of the 19th Psychology of Mathematics Education Conference* (Vol. 3. pp. 144-151). Sao Paulo, Brazil: PME Program Committee.

Cai, J. (2000). Understanding and representing the arithmetic averaging algorithm: An analysis and comparison of U.S. and Chinese students' responses. *International Journal of Mathematical Education in Science and Technology, 31*(6), 839-855.

Cai, J., & Gorowara, C. C. (2002). Teachers' conceptions and constructions of pedagogical representations in teaching arithmetic average. In B. Philips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics.* Cape Town: International Statistical Institute and International Association for Statistical Education.

Casey, S. A. (2008). *Subject matter knowledge for teaching statistical association* (Published doctoral thesis). Illinois State University. Retrieved from http://bit.ly/2jbnYJu

Cheah, U. H. (2001). *The construction of mathematical beliefs by trainee teachers' college: A multiple case study* (Unpublished doctoral thesis). Universiti Sains Malaysia.

Chew, C. M. (2007) *Form one students' learning of solid geometry in a phase-based instructional environment using the Geometer's Sketchpad* (Unpublished doctoral thesis). University of Malaya, Kuala Lumpur.

Chick, H. L., & Pierce, R. U. (2012). *Teaching for statistical literacy: Utilising affordances in real-world data*. *International Journal of Science and Mathematics Education*, *10*(2), 339-362.

Cobb, G. W., & Moore, D. S. (1997). Mathematics, statistics, and teaching. *American Mathematical Monthly*, 801-823.

Coffey, A., & Atkinson, P. (1996). *Making sense of qualitative data: Complementary research strategies*. Thousand Oaks, CA: Sage.

Cortina, J. L., Saldanha, L., & Thompson, P. (1999). Multiplicative conceptions of the arithmetic mean. In *Proceedings of the Twenty First Meeting of the North American Chapter of the International Group of the Psychology of Mathematics Education* (Vol. 2, pp. 466-472).

Creswell, J. W. (2008) Educational Research: Planning, Conducting, And Evaluating Quantitative and Qualitative Research.

Eichler, A., & Zapata-Cardona, L. (2016). Empirical Research in Statistics Education. In Empirical Research in Statistics Education (pp. 1-37). Springer International Publishing.

Elbaz, F. (1983). *Teacher thinking: A study of practical knowledge*, New York: Nichols.

Estrada, A., Batanero, C., & Fortuny, J. M. (2004). Un studio sobre conocimientos de estadìstica elemental de profesores en formación (Prospective teachers' knowledge on elementary statistics). *Educacón matemática*, *16*(1), 89-111.

Even, R. (1990). Subject matter knowledge for teaching and the case of functions. *Educational Studies in Mathematics, 21*(6), 521-544.

Even, R. (1993). Subject-matter knowledge and pedagogical content knowledge: Prospective secondary teachers and the function concept. *Journal for Research in Mathematics Education, 24*(2), 94-116.

Even, R., & Tirosh, D. (1995). Subject-matter knowledge and knowledge about students as sources of teacher presentations of the subject-matter. *Educational studies in mathematics*, *29*(1), 1-20.

Fatimah, S. (1997). *Skim penyelesaian masalah bagi guru matematik tingkatan dua* [*Problem solving schemes of form two mathematics teachers*] (Unpublished doctoral thesis). University of Malaya, Kuala Lumpur.

Fennema, E., & Franke, M. L. (1992). Teachers' knowledge and its impact. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 147-164). New York: Macmillan.

Frankenstein, M. (1998). Reading the World with Math: Goals for a Critical Mathematical Literacy Curriculum, In E. Lee, D. Menkart, & M. Okazawa-Rey (Eds.) *Beyond Heroes and Holidays: A Practical Guide to K-12 Anti-Racist, Multicultural Education and Staff Development*. Washington, DC: Network of Educators on the Americas.

Gal, I. (1994). *Assessment of interpretive skills*. Summary of working group, Conference Assessment Issues in Statistics Education, Philadelphia, PA.

Gal, I. (1995). Statistical tools and statistical literacy: The case of the average. *Teaching Statistics*, *17*(3), 97-99.

Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, *70* (1), 1–25.

Gal, I., Rothschild, K., & Wagner, D. A. (1990, April). Statistical concepts and statistical reasoning in school children: Convergence or divergence. In *annual meeting of the American Educational Research Association, Boston, MA*.

Garcia, C., & Garret, A. (2006). On average and open-ended questions. In A. Rosssman & B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics*. Salvador (Bahia), Brazil, International Statistical Institute and International Association for Statistical Education.

Garfield, J., Delmas, B., & Chance, B. (2003). The web-based ARTIST: Assessment resource for improving statistical thinking. In Symposium: Assessment of Statistical Reasoning to Enhance Educational Quality of AERA Annual Meeting.

Garfield, J., & Ben-Zvi, D. (2004). *The challenge of developing statistical literacy, reasoning and thinking*. Dordrecht, The Netherlands: Kluwer.

Garfield, J., & Ben-Zvi, D. (2007). How students learn statistics revisited: A current review of research on teaching and learning statistics. *International Statistical Review*, *75*(3), 372-396.

Gfeller, M. K., Niess, M. L., & Lederman, N. G. (1999). Pre-service teachers' use of multiple representations in solving arithmetic mean problems. *School Science and Mathematics*, *99*(5), 250-257.

Goldin, G. (2000) A scientific perspective on structures, task-based interviews in mathematics education research. In R. Lesh & AE. Kelly (Eds) *Research design in mathematics and science education* (pp 517–545). Erlbaum, Hillsdale.

Grossman, P. L., Wilson, S. M., & Shulman, L. S. (1989). Teachers of substance: Subject matter knowledge for teaching. *Knowledge base for the beginning teacher*, *27*.

Groth, R. E. (2007) Toward a Conceptualization of Statistical Knowledge for Teaching. *Journal for Research in Mathematics Education*, *38* (5), 427 -437.

Groth, R. E., & Bergner, J. A. (2006). Pre-service elementary teachers' conceptual and procedural knowledge of mean, median, and mode. *Mathematical Thinking and Learning*, *8*, 37-63.

Hill, H., & Ball, D. L. (2009). The curious—and crucial—case of mathematical knowledge for teaching. *Phi Delta Kappan*, *91*(2), 68-71.

Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education*, 372-400.

Hill, H. C., Schilling, S. G., & Ball, D. L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *The Elementary School Journal*, *105*(1), 11-30.

Hobden, S. (2014). When Statistical Literacy Really Matters: Understanding Published Information about the HIV/AIDS Epidemic in South Africa. *Statistics Education Research Journal, 13*(2), 72-82.

Hogan, T. P. (2007). *Educational Assessment: A practical introduction*. NJ: John Wiley & Sons, Inc.

Hunting, R. P. (1997). Clinical interview methods in mathematics education research and practice. *The Journal of Mathematical Behavior*, *16*(2), 145-165.

Hunting, R. P., & Doig, B. A. (1997). Clinical Assessment in Mathematics: Learning the Craft. Focus on learning problems in Mathematics, 19(3), 29-48.

Ministry of Education. (2006). *Integrated Curriculum for Primary Schools Curriculum Specifications Mathematics Year Four.* Kuala Lumpur, Malaysia: Author.

Ministry of Education. (2006). *Integrated Curriculum for Primary Schools Curriculum Specifications Mathematics Year Five.* Kuala Lumpur, Malaysia: Author.

Ministry of Education. (2006). *Integrated Curriculum for Primary Schools Curriculum Specifications Mathematics Year Six.* Kuala Lumpur, Malaysia: Author.

Ministry of Education. (2006). *Integrated Curriculum for Primary Schools Curriculum Specifications Mathematics Form Three.* Kuala Lumpur, Malaysia: Author.

Ministry of Education. (2006*). Integrated Curriculum for Primary Schools Curriculum Specifications Mathematics Form Four.* Kuala Lumpur, Malaysia: Author.

Ministry of Education. (2006). *Integrated Curriculum for Primary Schools Curriculum Specifications Additional Mathematics Form Four.* Kuala Lumpur, Malaysia: Author.

Ministry of Education. (2006). *Integrated Curriculum for Primary Schools Curriculum Specifications Additional Mathematics Form Five.* Kuala Lumpur, Malaysia: Author.

Jacobbe, T. (2007). *Elementary school teachers' understanding of essential topics in statistics and the influence of assessment instruments and a reform curriculum upon their understanding* (Doctoral Thesis).

Jacobbe, T. (2012). Elementary school teachers' understanding of mean and median. International Journal of Science and Mathematics Education, 10(5), 1143-1161.

Jacobbe, T., & Carvalho, C. (2011). Teachers' understanding of averages. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching Statistics in School Mathematics-Challenges for Teaching and Teacher education: A Joint ICMI/IASE Study*.

Jairaman, K., Akmar, S. N., Zamri, S., Rahim, A., & Syrene, S. (2016a). A Pre-Service Mathematics Teacher's Subject Matter Knowledge of the Mode: A Case Study. Malaysian Online Journal of Educational Sciences, 4(3), 1-11.

Jairaman, K., Akmar, S. N., Zamri, S., Rahim, A., & Syrene, S. (2016b). Pre-Service Mathematics Teachers' Subject Matter Knowledge of the Mode as Measures of Central Tendency. Jurnal Kurikulum dan Pengajaran Asia Pasifik, 4(3), 11-19.

Jairaman, K., Akmar, S. N., Zamri, S., Rahim, A., & Syrene, S. (2015). Pre-Service Mathematics Teachers' Subject Matter Knowledge of Average. Sixth International Conference on Science and Mathematics Education CoSMEd Penang, Malaysia.

Koe, C. A. (1992, December). *Pencapaian matematik guru pelatih di maktab perguruan* [*Trainee teachers' mathematics achievement in teacher training colleges*]. Paper presented at the Persidangan Kebangsaan Peningkatan Pengajaran Pembelajaran Matematik di Institut/Maktab Perguruan Malaysia [National Conference on the Improvement of Mathematics Teaching and Learning at the Institute/Teacher Training Colleges Malaysia]. Melaka, Malaysia.

Koleza, E., & Kontogianni, A. (2013, January). Assessing statistical literacy: What do freshman know. In *Proceedings of the 8th Congress of the European Society for Research in Mathematics Education, CERME* (Vol. 8).

Konold, C., & Garfield, J. (1992). *Statistical reasoning assessment: Part 1. Intuitive Thinking*. Amherst, MA: Scientific Reasoning Research Institute, University of Massachusetts.

Konold, C., & Pollatsek, A. (2002). Data analysis as the search for signals in noisy processes. *Journal for Research in Mathematics Education*, 259-289.

Konold, C., & Pollatsek, A. (2004). Conceptualizing an average as a stable feature of a noisy process. In *The challenge of developing statistical literacy, reasoning and thinking* (pp. 169-199). Springer Netherlands.

Krefting, L. (1991). Rigor in qualitative research: The assessment of trustworthiness. *American Journal of Occupational Therapy, 45*, 214-222.

Lampert, M.: 1988, The teacher's role in reinventing the meaning of mathematical knowing in the classroom, In M. J. Behr, C. B. Lacampagne, & M. M. Wheeler (Eds.), *Proceedings of the 10th Annual Meeting of PME-NA*, *DeKalb*, Ill., (pp. 433-480).

Learning Mathematics for Teaching (2006). *A coding rubric for measuring the mathematical quality of instruction* (Technical Report LMT1.06). Ann Arbor, MI: University of Michigan, School of Education.

Leavy, A., & O'Loughlin, N. (2006). Pre-service teachers understanding of the mean: Moving beyond the arithmetic average. *Journal of Mathematics Teacher Education, 9*, 53-90.

Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Mahwah, NJ: Lawrence Erlbaum Associates.

Maher, C. A., & Sigley, R. (2014). Task-based interviews in mathematics education. In *Encyclopedia of Mathematics Education* (pp. 579-582). Springer Netherlands.

Maxwell, J. A. (2005). *Qualitative research design: An interactive approach* (2nd ed.) Thousand Oaks, CA: Sage.

McDiarmid, G. W. (1989). Why Staying One Chapter Ahead Doesn't Really Work: Subject-Specific Pedagogy. Issue Paper 88-6.

McGatha, M., Cobb, P., & McClain, K. (2002). An analysis of students' initial statistical understandings: Developing a conjectured learning trajectory. *The Journal of Mathematical Behavior*, *21*(3), 339-355.

McMillan, J. H. (2001). *Classroom assessment: Principle and practice for effective instruction* (2nd ed.). Boston: Allyn and Bacon.

Merriam, S. B. (1998). *Qualitative Research and Case Study Applications in Education. Revised and Expanded from" Case Study Research in Education"*. San Francisco, CA: Jossey-Bass Publishers.

Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation: Revised and expanded from qualitative research and case study applications in education.* San Franscisco: Jossey-Bass.

Mevarech, Z. R. (1983). A deep structure model of students' statistical misconceptions. *Educational Studies in Mathematics*, 14, 415-429.

Mokros, J., & Russell, S. J. (1995). Children's concepts of average and representativeness. *Journal for Research in Mathematics Education*, 26, 20-39.

Mosenthal, P. B., & Kirsch, I. S. (1998). A new measure for assessing document complexity: The PMOSE/IKIRSCH document readability formula. *Journal of Adolescent and Adult Literacy*, *41*(8), 638-657.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Aurthor.

Ng, S. F. (1995). *Malaysian pre-service primary mathematics teachers and their lecturers: Practice and beliefs about mathematics, teaching and learning* (Unpublished doctoral thesis). University of Birmingham.

Nik Azis, N. P. (1987). *Children's fractional schemes* (Unpublished doctoral thesis). University of Georgia, Athens, GA.

Nik Azis, N. P. (1995). *Penghayatan matematik KBSR dan KBSM: Agenda tindakan* (Edisi kedua) [*Appreciation of the integrated curriculum of primary and secondary school mathematics: Action agenda (2nd ed.)*]. Kuala Lumpur: Dewan Bahasa dan Pustaka.

Nik Azis, N. P. (1999). *Pendekatan konstruktivisme radikal dalam pendidikan matematik* [*Radical constructivism approaches in mathematics education*]. Kuala Lumpur: Penerbit Universiti Malaya.

Nik Azis, N. P. (2003, October). *Mathematics education in the 21st century: Meeting the challenges of a changing world*. Paper presented at the International Conference on Science and Mathematics Education: Which Way Now? University of Malaya, Kuala Lumpur.

Nik Azis, N. P. (2008). *Isu-isu kritikal dalam pendidikan matematik* [*critical issues in mathematics education*]. Kuala Lumpur: Penerbit Universiti Malaya.

Nik Azis, N. P., & Ng, S. N. (1990). Pelaksanaan matematik KBSM: Fenomena tahun pertama [Implementation of the Malaysian Integrated Curriculum of Secondary Mathematics: First year phenomena]. *Masalah Pendidikan* [*Issues in Education*], *14*, 31-45.

North, D., Gal, I., & Zewotir, T. (2014). Building capacity for developing statistical literacy in a developing country: lessons learned from an intervention. *Statistics Education Research Journal, 13*(2), 15-27.

Olfos, R., & Estrella, S. (2010). Chilean primary teachers challenged to build PCK for statistics. In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8) Ljubljana, Slovenia*. Voorburg, The Netherlands: International Statistical Institute.

Patton, M. Q. (1990). *Qualitative evaluation and research methods*. SAGE Publications Inc.

Pollatsek, A., Lima, S., & Well, A. D. (1981). Concept or Computation: Students' understanding of the mean. *Educational Studies in Mathematics*, 12, 191-204.

Popham, W. J. (2000). *Modern educational measurement: Practice guidelines for educational leaders* (3rd Ed.). Boston: Allyn and Bacon.

Preece, R., & Preece, R. A. (1994). *Starting research: an introduction to academic research and dissertation writing*. London: Pinter.

Russell, S. J., & Mokros, J. R. (1991). What's typical? Children's ideas about average. In D. Vere-Jones (Eds.), *Proceedings of the Third International Conference on Teaching Statistics* (pp. 307-3130). Voorburg, Netherlands: International Statistical Institute.

Santos, R., & da Ponte, J. P. (2013). Prospective elementary school teachers' interpretation of central tendency measures during a statistical investigation. In Proceedings of the Eighth Congress of the European Society for Research in Mathematics Education (CERME8). Ankara: Middle East Technical University.

Sharifah Norul Akmar, S. Z. (1997). *Skim penolakan integer pelajar tingkatan dua* [*Integers subtraction schemes of form two students*] (Unpublished doctoral thesis). University of Malaya, Kuala Lumpur.

Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 957-1010). Greenwich, CT: Information Age Publishing and National Council of Teachers of Mathematics.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14.

Silver, E. A. (1986). Using conceptual and procedural knowledge: A focus on relationships. In J. Hiebert (ed.), *Conceptual and Procedural Knowledge: The Case of Mathematics* (pp. 181-198). New Jersey: Lawrence Erlbaum Associates, Inc.

Sirnik, M., & Kmetič, S. (2010, July). Understanding of arithmetic mean. In *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8*).

Strauss, S. & Bichler, E. (1988). The development of children's concepts of the arithmetic average. *Journal for Research in Mathematics Education*, *19*(1), 64-80.

Thames, M., Sleep, L., Bass, H., & Ball, D. L. (2008, July). Mathematical knowledge for teaching (K-8): Empirical, theoretical, and practical foundations. In *Paper presented at the 11th International Conference on Mathematics Education (ICME-11) for Topic Study Group* (Vol. 27).

Thompson, A. G. (1984). The relationship of teachers' conceptions of mathematics and mathematics teaching to instructional practice. *Educational Studies in Mathematics, 15*, 105- 127.

Vinner, S. (1983). Concept definition, concept image and the notion of function. *International Journal of Mathematical Education in Science and Technology*, *14*, 239-305.

Wallman, K. K. (1993). Enhancing statistical literacy: Enriching our society. *Journal of the American Statistical Association*, *88* (421), 1-8.

Wanta, W. (1997). *The public and the national agenda: How people learn about important issues*. Mahwah, NJ: Lawrence Erlbaum.

Watson, J.M. (1997). Assessing statistical literacy through the use of media surveys. In I. Gal & J. Garfields (Eds), *The assessment challenge in statistics education* (pp. 107 – 121). Amsterdam, the Netherlands: The International Statistical Institute/ IOS Press.

Watson, J.M. (2006). *Statistical Literacy at School: Growth and Goals*. Lawrence Erlbaum, New Jersey, USA.

Watson, J.M., & Callingham, R.A. (2003). Statistical literacy: A complex hierarchical construct. *Statistics Education Research Journal*, *2*(2), 3-46.

Watson, J.M., & Callingham, R.A. (2013). PCK and Average, In V. Steinle, L. Ball & C. Bardini (Eds.), Mathematics Education: Yesterday, today and tomorrow. *Proceedings of the 36th annual conference of the Mathematics Education Research Group of Australasia*. Melbourne, VIC: MERGA.

Watson, J. M., & Chick, H.L. (2012). Average revisited in context. *Proceedings of the 35th Annual Conference of the Mathematics Education Research Group of Australasia* (pp. 753-760). Mathematics Education: Expanding Horizons.

Watson, J. M., & Moritz, J. B. (1999). The development of concepts of average. *Focus on Learning Problems in Mathematics*, *21*(4), 15-39.

Watson, J. M., & Moritz, J. B. (2000). The longitudinal development of understanding of average. *Mathematical Thinking and Learning*, *2*(1&2), 11-50.

Watson, J. M., & Moritz, J. B. (2000). Development of understanding of sampling for statistical literacy. *The Journal of Mathematical Behavior*, *19*(1), 109-122.

Watson, J. M. (2011). Foundations for improving statistical literacy. *Statistical Journal of the IAOS 27*, 197–204.

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, *67*(3), 223-248.

Wun, T. H. (2010). *Pre-service mathematics teachers' subject matter knowledge of perimeter and area* (Unpublished doctoral thesis). University of Malaya, Kuala Lumpur.

Yin, R. K. (2003). *Case study research* (Vol. 5). Thousand Oaks, California.

List of publications

1. Jairaman, K., Akmar, S. N., Zamri, S., Rahim, A., & Syrene, S. (2016). A Pre-Service Mathematics Teacher's Subject Matter Knowledge of the Mode: A Case Study. *Malaysian Online Journal of Educational Sciences*, *4*(3), 1-11.
2. Jairaman, K., Akmar, S. N., Zamri, S., Rahim, A., & Syrene, S. (2016). Pre-Service Mathematics Teachers' Subject Matter Knowledge of the Mode as Measures of Central Tendency. *Jurnal Kurikulum dan Pengajaran Asia Pasifik, 4*(3), 11-19.
3. Jairaman, K., Akmar, S. N., Zamri, S., Rahim, A., & Syrene, S. (2015). Pre-Service Mathematics Teachers' Subject Matter Knowledge of Average. *Sixth International Conference on Science and Mathematics Education* CoSMEd Penang, Malaysia.